

Too much fall-out from summit

Last week's summit meeting in Bonn was no more of a failure than its predecessors. These meetings do not settle issues, but have the virtue of putting them on the international agenda.

TAKE seven heads of government, each preoccupied with urgent domestic business, and throw them together for a couple of days with instructions to consider the world's urgent problems from the other fellow's point of view. It is a daunting assignment. Usually, after the event, the public relations people are hard pressed to conceal the disturbing truth that even world statesmen, as they are called, do not rise to the grand occasions. Their simplest trick is to make public the quarrels there will certainly have been on relatively minor issues, but at the same time to suggest that none of this can presage permanent disunity (which is usually correct). So what can the statesmen claim they have accomplished when they return to face their own electors? Prudently, they arrange that each among them has some cheerful souvenir to take home. On this occasion, there was lukewarm support for President François Mitterrand's plan that high technology should be used to help Africa feed itself, general approval almost indistinguishable from indifference for the British government's economic policies and muted acquiescence in US plans for star wars research. But the statesmen at Bonn failed to agree on the US demand that there should be a further round of talks next year on the liberalization of international trade and on the French demand that international monetary reform should be a precondition of such talks.

The temptation to blame the French for the failure at Bonn should be suppressed, if only because the French, and especially President Mitterrand, have consistently done their best to tackle other than immediate problems. Four years ago, at Fontainebleau, the cry was that technology should be made the instrument of renewed economic growth. This year, at Bonn, the plan was to rid Africa of famine by an engineered green revolution coupled, wherever appropriate, with the use of Earth resources satellites and the like. President Mitterrand is entirely right to believe that technology can work wonders, and would now be within his rights to be irritated that his fellow-statesmen do not share his vision. Where he goes wrong is to suggest that technology, admittedly a necessary part of the solution of these problems, is for that reason also sufficient. That is the illusion. While the climate and the soil of much of developing Africa are inimical to agriculture, the reasons why food production is inadequate have more to do with the willingness of African governments to sacrifice the prosperity of their peasant and farming populations to the interests of their urban voters than with the physical conditions under which crops must be grown. A summit of the seven will not put that to rights.

The French demand for monetary reform is similarly stronger on good intentions than on realism. Rapid currency fluctuations are unsettling, and bad for world trade (which may be true), so let us legislate against them. That seems to be the argument. It is a classic case of therapy by the removal of symptoms. But President Mitterrand is well placed to know that before fixed currency rates were finally abandoned in 1972, international financial crises happened just as often, but took a different form. The immediate cause of the troubles of the past few months is the large budget deficit of the federal government of the United States, and the way in which this is being financed (by borrowing from abroad rather than by printing domestically inflationary dollar bills). Since there is no way in which sovereign governments can be prevented by their trading partners from running budget

deficits or even, if they choose, from letting domestic inflation run riot, the only present basis for monetary reform is an agreement between the summit states to limit the effects of bad budgeting on the international money markets. If Mr. Mitterrand had been pressing for true reform along such lines, rather than for a return of the bad old days, the United States would probably still have disagreed with him, but then he would have had reason on his side.

The US demand for a further round of talks on international trade beneath the umbrella of the General Agreement on Tariffs and Trade (GATT) is a more complicated issue. That GATT should exist is a sign of grace. That the United States should have taken the initiative in pressing for an extension of the free-trade regime is decidedly another. Politically, it might have been easier for the US administration to give in to the recent demands in Congress for a greater measure of protection from imports from overseas. But the United States was asking last weekend for an extension of the rules governing international trade that would strike at the roots of many shaky institutions elsewhere. The proposal that the rules on free trade should be extended to agricultural produce might yet make a market for Texas beef in Japan, but would certainly spell trouble for the practice of the European Community of subsidizing farmers by means of artificially high prices for food in domestic markets, while retaining the right to dump the surplus on the international markets. (One of last weekend's little ironies is that, while Chancellor Helmut Kohl was seeking emollient ways of making the Bonn communique seem good, his agriculture minister had dug in his heels, on behalf of West German farmers, against a proposal that the price of European cereals should be reduced.) The truth is that, on free trade, the United States has reason on its side and should be helped to have its way.

Star wars, the other agenda item, is more of an embarrassment for all. Two years ago, when the Strategic Defense Initiative, as it was later called, took much of the US administration by surprise, most other governments were either sceptical (as in Western Europe) or hostile (the Soviet Union). With the passage of time, the administration seems to have calculated that doubters might be turned into supporters if they were offered participation, but the terms on which this will be possible are not clear. But most of the summit governments, and certainly their electors, would be happier if the whole scheme had never been thought of. Some of them will talk about participation in the hope that helping will give them influence, but are alarmed that star wars may jeopardize the missile talks under way at Geneva.

Agreement on that issue, like most of the others raised in Bonn, lies in the future. To next year's summit, perhaps? Many of the issues raised in Bonn cannot wait that long. The US deficit will either have been made to go away, or there will be a different set of economic problems with which to grapple. But President Mitterrand may find, by next year, that he has been luckier than he now thinks likely on monetary reform, given the way in which bankers have taken fright over recent developments. If star wars have confounded the Geneva negotiations, or if these have broken down for other reasons, there will be more urgent things to talk about, which only goes to show that these summits, whatever they fail to accomplish, have become a valuable agenda of our anxieties. □



Running telescopes

Managing British astronomy is up for talks. Here is a recipe for the committee responsible.

THE first tangible product of the review of the pattern of British academically related research to which the Science and Engineering Research Council committed itself six months ago is hardly the revolution some had feared (or, sometimes, hoped for), but is none the less interesting on that account. The nub of the decision about research in astronomy announced last week is for the time being unremarkable, even hackneyed: there is to be another committee, the composition of which will be decided only next week. But in the course of inviting comments, the council says its objective is to decide what arrangements there should be for managing the two modern observatories in which the British have a majority share, at La Palma and in Hawaii. (The Netherlands, Spain, Denmark and Sweden are differently involved in one or other or even both of the projects.) By putting the question in those terms, the council has side-stepped what in the past has seemed the central question in British astronomy, that of whether the two observatories (at Herstmonceux and Edinburgh) should be merged and, if so, on which site. To ask instead the question how the new observatories should ideally be managed should provide a more constructive agenda for discussion than the many earlier enquiries on related subjects.

As things are, the Royal Greenwich Observatory (at Herstmonceux) is responsible for the La Palma observatory, to which it has donated its largest telescope, the 92-inch Isaac Newton reflector. Edinburgh, on the other hand, has had operational responsibility for the Anglo-Australian telescope, roughly 18,000 kilometres distant. (Nothing is said in last week's announcement of the importance of that instrument in the future scheme of things, confirming earlier guesses that the research council plans to trade its share of the capital equipment for a measure of observing time.) But yet another laboratory, the Rutherford Appleton Laboratory in the middle of southern England, has an important place in the management of British astronomy, both in the operations of the computer network by which data are exchanged between university laboratories and in supervision of the millimetre-wave telescope being built on Mauna Kea. On the face of things, British astronomers are over-well catered for in these respects. What the new committee should be asking is whether they should now be given a chance to look after themselves.

There is now plenty of recent experience to suggest how this should be arranged. Like other kinds of laboratories, observatories need directors, preferably persons of distinction in the field concerned, astronomy in this case. Even when the chief users are from universities, observatories also need to have a resident scientific and technical staff. Given the geography of the new British ventures, these will presumably be chosen to be as small as possible. For the rest, there need to be oversight committees, essentially trustees, and committees for the allocation of observing time, which again need not always meet in the shadow of the instruments they service. To give this pattern of management conviction with the users, some way should be found of incorporating users or their representatives in these managerial functions. These are the essential ingredients of the management structure, which happen not to involve the existing observatories at Herstmonceux and Edinburgh.

But this, of course, is not the whole of the fabric of support for a lively astronomy enterprise. The development of instruments has been shown in recent years to have been crucial to the improvement of techniques, at both telescopes, while the management and storage of data is likely to become just as important in the years ahead. But the people best placed to know what instruments will yield worthwhile results are potential users and their frequent attendants, the people with a zeal (and flair) for making things a little more sensitive or accurate, from which it follows that the research council should make grants to university astronomers for the development of instruments (which

already happens), keeping up its sleeve some means for building what the enthusiasts design (best done, alongside the development of instruments for satellites, at the Rutherford Appleton Laboratory). Data management needs to be done by astronomers, not toilers in some huge laboratory, which argues for transferring this responsibility to one of the two observatories in the United Kingdom. With its recent success in designing data handling techniques, Edinburgh seems a better bet than Herstmonceux.

This is the kind of pattern to which logic leads — a single smaller observatory based in Britain, more responsibility (as for instrument design) given to the university users of the new equipment and central support in engineering from a large laboratory. So why bother with a committee? Because committees are about politics as well as reason. The choice of which observatory to close is bound to be invidious, and so is most easily made by people with an appearance of detachment. That is how the argument goes. So the research council ultimately in charge should simplify the committee's work by making clear in advance how it would deal with the people, possibly counted in hundreds, whose jobs would vanish under a brisk plan for reorganization. And the council should do its level best to practise what it has been preaching in the past few years, that astronomy is not only an end in itself but also a way of acquiring skills of general applicability, by shouldering the task of finding acceptable and useful jobs for those who may be displaced. At a time when Britain, against the recent trends, is about to be better blessed with facilities for observation in astronomy, it would be absurd that the dole queues should be lengthened by an influx of professionals in the field.

Beware junk bonds

Business in the United States, always inventive, seems bent on turning fiction into truth.

THE board-game "Monopoly" was designed in Atlantic City to mimic the real world of property speculation. Now, the financial wizards of Wall Street and similar places seem bent on adapting "Monopoly" to real life. Shadowy figures with colourful names such as T. Boone Pickens have hit on a novel way of buying up desirable companies in the stock market. The stratagem is ingenious and simple. A publicly-quoted company is singled out as a target, often on the strength of its diversity and the likelihood that its business assets are undervalued. Then a second corporation is created which offers to buy a proportion of the stock in the target in return for securities of its own, usually in the form of interest-bearing certificates. To make the deal attractive to the shareholders, the rate of interest is fixed at a level great enough to promise a return greater than the dividends earned from the target company. To lend reality to the offer, the company usually has to spend cash buying a proportion of the outstanding stock in the markets. Under United States law, it will suffice for the raider's purpose to acquire just that proportion of stock to gain control of the target company.

That is phase one. The problem now is that the shell company that has won control of its target is saddled with a load of debt, which can only be discharged by selling off the assets of what was originally the target company, exchanging the cash proceeds for the outstanding debt. In principle, these schemes will work so long as the assets can be sold for more than the nominal cost of buying half the target company. In practice it is often more profitable to offer some cash with the original purchase, for then the purchase price may be much lower. One way and another, the junk bonds outstanding from the recent frenzy of takeovers in the United States continue to be serviced. So far as anybody can tell, none of the schemes financed by junk bonds has yet come unstuck. But the amounts of money involved are so large (\$9,000 million for Gulf Oil) and the basis of these deals so shadowy, that some of them are bound at some stage to run into trouble, perhaps because the assets are no longer as saleable as they seemed. Then a lot of people who thought themselves rich will find themselves poor, the common experience after the Great Crash of the 1930s from which "Monopoly" sprang. □

US agricultural research

Mission-agency seeks relief from Congress

Washington

THE Agricultural Research Service (ARS), the US Department of Agriculture's sprawling intramural research agency, is emerging from three years of austerity (some self-imposed) and reorganization with a mixed success in its efforts to polish its tarnished image with the scientific community.

ARS's administrator, Dr Terry Kinney, has scored one major success by reducing administrative overheads by \$12 million out of the agency's \$487 million annual budget. Less successful have been Kinney's attempts to challenge Congress's insatiable appetite for building new laboratories that for the most part serve only parochial interests.

And the most long-standing criticism of ARS — that it is not keeping pace with new developments in basic research, especially molecular genetics — persists, in spite of one notable ARS initiative, the creation last year of a plant gene-expression centre in conjunction with the University of California.

That criticism was reiterated most recently by a National Academy of Sciences panel which scrutinized ARS's research programmes at the request of ARS itself. The panel's report again emphasized the need to concentrate on developing basic understanding of plant physiology and development, insect neurobiology, immune responses in animals and other fundamental molecular processes in domestic plants and animals as well as pests.

Standing in contrast to the panel's recommendations is ARS's own 1984 annual report of its research achievements, which highlights applications almost exclusively — a new rotary disk for harvesting soya beans, a pelletized guayule seed that improves germination, a pilot plant for automatic tanning of cattle hides, a new kind of spaghetti made with 10 per cent bran.

The academy panel, chaired by Ralph Hardy, recently head of life sciences at DuPont and now president of BioTechnica, also recommended some drastic steps for ARS if it is to create the proper "climate" for good research. For example, the panel said ARS needs to be a lot tougher about offering permanent positions to scientists.

At present, the panel said, virtually all ARS scientists who are reviewed at the end of a one-year probationary period are promoted to permanent staff positions; the probationary period should be extended to five years and candidates should be evaluated by an outside group of experts.

The panel also suggested a substantial increase in ARS's postdoctoral programme;

instead of 25 postdocs (actually, ARS officials say, the true figure is closer to 275), there should be at least 750. And instead of devoting 90 per cent of research budgets to salaries — as at some of the ARS's 147 laboratories and research centres — salaries should be at most 75 per cent, and in some cases as low as 60 per cent, the remainder going towards properly equipping the laboratories.

The 147 laboratories are a problem in themselves. The panel noted, as have many others before, that that is simply too many. Research groups are fragmented, preventing the formation of a "critical mass" of scientists needed to make progress.

ARS officials say that they have already dealt with many of these complaints, but that their hands are tied on the others. Federal personnel rules, for example, prohibit use of outside consultants in hiring and promotion decisions. But ARS is now negotiating with the federal Office of Personnel Management to lengthen the current one-year probationary period to three years.

Congress has been the greatest obstacle to ARS's attempts to shut down centres and laboratories that have outlived their usefulness, or which fragment the research effort. ARS did succeed in transferring the staff of the fire-ant research laboratory in Gulf Port, Mississippi, to its Gainesville, Florida, insect research centre. It hopes to consolidate several other laboratories this year.

But Congress has effectively turned down a request to rescind funds appropriated in fiscal years 1984 and 1985 for construction at eight facilities (including a

Forage Seed Production and Research Center and the National Soil Tilth Center). Congress has also continued to create new facilities, such as a national centre for leptospirosis in Ames, Iowa, and a Children's Nutrition Research Center in Houston, Texas.

One way that ARS hopes to avoid being saddled with still more laboratories in the future is by writing new language into the farm bill, which is up for renewal this year. As things stand, the only way for a state to obtain federal money for a research facility is by having ARS build and operate it. The new language would open an escape valve, allowing Congress to appropriate matching funds for states to build their own facilities when they consider this necessary.

Dr Mary Carter, ARS's associate director, says that the agency has already taken steps to meet the panel's criticism of high personnel costs, and that many laboratories are already down to 65 or 70 per cent. Some laboratories where the figure is deemed too high have been ordered not to fill vacancies to correct the problem.

Carter also said that ARS is committed to expanding its programmes of postdoctoral research. Besides the 50 positions that are being created directly by the administrator's office, the research laboratories themselves hire some 250 or so postdocs.

But ARS, always under pressure from Congress — particularly Representative Jamie Whitten (Democrat, Mississippi), chairman of the Appropriations Committee — to "do something for the farmer", is drawing the line at any major tilt in favour of basic research. "We're a mission-oriented agency," Carter said, emphasizing the new ARS party line that biotechnology is merely a group of "methodologies" that can be put to use at all of its laboratories. ARS has explicitly rejected the notion of establishing biotechnology or basic research centres.

Stephen Budiansky

Land bows out from Polaroid

EDWIN Land, the 76-year-old founder of the Polaroid Corporation, has broken his last tie with the company. Last week, Land announced he would sell his remaining shares, which account for 8.3 per cent of Polaroid stock, and turn over a sizeable portion of the proceeds to the non-profit Rowland Institute for Science, a vaguely utopian haven for basic research that Land set up in 1981 by selling off \$38 million worth of stock.

Land's remaining stock has an estimated market value of \$71 million, \$28 million of which is to go to the Rowland Institute.

Land resigned as chairman of Polaroid in 1982 (see *Nature* 298, 701; 1982) during a rough period for the company. The company's Polavision instant movie film, which hit the market in 1977, just as home

video equipment was becoming available, proved a colossal flop, costing the company a reported \$68 million. Although Polaroid's earnings began to recover from that setback in 1983, the company hit rough times again in the third quarter of last year.

The Rowland Institute, along with the Rowland Foundation that Land established in 1960 to support research and education, has increasingly been his consuming interest in recent years.

The institute, with a staff of about 80 and a new building on the Charles River in Cambridge, Massachusetts, was described by Land at the time of its founding as a place where scientists could return to "an older way of doing science", without large research teams or big machines.

Stephen Budiansky

France

New scourge for universities

In a rather extraordinary and for the universities somewhat disturbing development in French higher education, a principal critic of the French university system has been appointed president of a new universities "evaluation committee".

The man in question, M. Laurent Schwartz, is a highly-respected mathematician at the Ecole Polytechnique, one of the "ivy-league" *grandes écoles* of France. But he is also the author of two influential diatribes attacking the French universities and prescribing certain medicines, such as a greater dose of selection and élitism, which do not appeal to the majority of university staff. These were appointed in the egalitarian boom years after the 1968 student rebellion when the universities both shifted well to the political left and quadrupled in size. To them, a Schwartzian new broom is a fearsome thing.

But how much power will M. Schwartz have? This will not be clear until his new committee, the *Comité National d'Evaluation des Universités* (CNEU), has been in place for a year and delivered its first report. But that this may come around the time of next year's general elections in France, when, according to present opinion polls, the right will return to government, will cause further alarm.

In principle CNEU is being presented as equivalent to Britain's University Grants Committee (UGC), which distributes British universities' annual "recurrent grant" from the Department of Education and Science to cover salaries and basic university expenditure such as heating and building costs. But there are important differences. On the whole, UGC is composed of university heads of department representing a cross-section of academic disciplines. In the face of university cuts, therefore, it has tended to act as a defensive body, the only notable exception being when UGC cut grants to certain smaller, mostly technological universities on the debatable basis that these universities attracted students with poorer A-level grades than others such as Oxford and Cambridge. This, in a sense, was passing the buck.

But the French CNEU promises to be different. Apart from M. Schwartz, the nine members of the "academic" side of the committee will include physicist M. Pierre Aigrain, former science minister under the previous (and conservative) French President, Giscard d'Estaing, alongside other representatives of the *grandes écoles*, and few from the university system itself. The committee is completed by five members of what might be described as an "industrial" side of the committee. Altogether, then, CNEU cannot be described as representative of the universities, as can UGC. Rather, CNEU represents a new structure on top of the universities: a new level of control.

The appointments to CNEU were announced only last week. But its terms of reference were defined in a higher education bill last year. According to these, CNEU will regularly review the work funding of the universities, considering both their research and teaching. CNEU will also measure each university against its multi-annual plan — under which a university

agrees development plans with the ministry of education in exchange for a secure 3–5 year budget.

The committee will also review the new PhD system, which began only this year, and will not report to the minister of education but to the Elysee itself: to the President of the Republic, presently M. François Mitterrand.

It seems possible that M. Mitterrand, new scourge of the universities, while entering from the left, will end up by leaving from well to the right. **Robert Walgate**

NIH

More competition, more disquiet

Washington

THE Reagan administration's plan to restrict to 5,000 the number of extramural research project grants awarded by the National Institutes of Health (NIH) is having a marked effect on the priority score required of successful applicants. The estimated aggregate cut-off point for 1985 is 170, although because there have been more high-quality applications than expected, the figure may turn out to be as low as 160. The payline, as the cut-off is called, varies between institutes from less than 150 to over 200.

The Congress had requested 6,500 grants for 1985; if it prevails in the argument with the administration, the payline would be around 190. According to the latest rumours, a compromise of around 5,800 grants is likely.

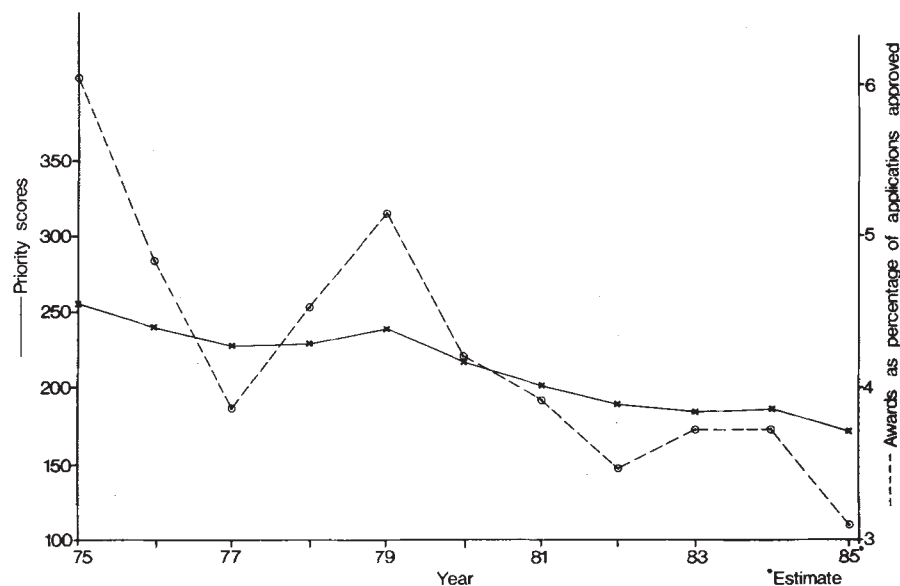
Meanwhile, intramural research at NIH is also facing problems. The administration's proposed budget for NIH intramural research in 1986 is \$561 million, down \$16 million from 1985. NIH officials estimate

that \$619 million would be needed in 1986 to maintain research at its present level.

In addition, the institutes as a whole have to lose 450 man-years from their staff by 30 September, with a reduction of a further 150 man-years during the following year. The number of staff that will have to be lost by September will probably be over 600, because the instruction to reduce staff (from the Office of Management and Budget) was received only in January, one-third of the way into the fiscal year.

The problem is further exacerbated in those institutes already employing more staff than the Office of Management and Budget had assumed. Institutes in this category have initiated a "semi-freeze" policy, whereby two jobs must be lost before a single replacement can be hired. NIH officials are concerned that young untenured researchers are being frozen out and that administrative and domestic staff, who are usually tenured, will as a consequence be a growing proportion of the whole.

Tim Beardsley



The broken line shows the percentage of grant applications approved by NIH study sections that were actually funded; the solid line shows the so-called "payline" — the priority score that 90 per cent of the funded applications had to better in order to receive funding. Priority scores range from 100 (excellent) to 500 (poor). The 1985 estimates assume 5000 grants this year; the actual number is still under negotiation.

Creation science

Rot sets in in Queensland

St Lucia, Queensland

IN Australia's sunshine state of Queensland, worried academics and teachers are preparing for a re-enactment of the "creation science" debate of the past few years in the United States. Their concern is that Queensland's Minister for Education, Mr Lin Powell, has ordered the teaching of creationism in state high schools. Whenever teachers raise the subject of evolution, they are also now obliged to present the conflicting arguments for creation and catastrophism.

Hitherto, many of Queensland's teachers have exploited a convenient loophole in Mr Powell's edict: they were not instructed how to teach creation science. But the minister now intends to close that loophole by spelling out to teachers exactly how they should present creationism in the classroom. His plans are being formulated with advice from the Creation Science Foundation (CSF), an organization that maintains a constant barrage of creationist propaganda from its headquarters in the Brisbane suburb of Sunnybank. At present, Queensland is the only Australian state where creationism is officially included in the school science curriculum, although some academics and teachers believe that it is being taught unofficially elsewhere.

CSF, which describes itself as a "faith-funded organization", publishes the glossy magazine *Ex Nihilo* at an annual cost of more than \$40,000. The magazine is delivered to all state high schools in Queensland (with approval from the Minister for Education), and also finds its way into school libraries in other states.

Late last year, the librarian at a high school in the Northern Territory judged *Ex Nihilo* to be unsuitable for children and requested CSF to send no further issues. CSF gave a prompt response in its newspaper *Prayer News* — it published the name and address of the offending school and urged its readers to write expressing disapproval. The editor of *Prayer News* commented that it was a "dangerous situation" when a library refused to accept a gift of *Ex Nihilo*. Eventually, the hapless school librarian capitulated under mounting pressure from CSF and its supporters.

While *Ex Nihilo* brings creationism to a general audience (including children), CSF has also launched a second periodical destined for only the most erudite creationists. The *Ex Nihilo Technical Journal* is laid out in the format of many reputable scientific journals and purports to give an in-depth treatment of the latest developments in creation research.

CSF, not content merely to make inroads into the educational system, now also offers its own correspondence courses in creation science, which may be attractive to teachers seeking additional "qualifications". The gratifyingly large number of

enrolments (more than 90) has prompted CSF to start selling its courses in the United States, where it already has representatives to promote sales of its literature. Most recently, CSF has revealed plans for a "world class museum as a teaching centre on creation".

So far, the creationist campaign has met no organized opposition from the scientific community. Not a single scientific society has spoken out against the bogus science of creationism. The Australian and New Zealand Association for the Advancement of Science has remained silent, as has the Australian Academy of Sciences. Representatives from all major scientific organizations, gathered in Canberra on 16 April for the inaugural meeting of the National Committee for the Promotion of Science

and Technology, might well have been expected to voice some opinion on the issue, but did not.

Nevertheless, many individuals — scientists and teachers, students and parents — are beginning to express concern about the unchecked growth of creationism. On 22 March, an audience of nearly 400 gathered at the University of Queensland to hear a series of lectures exposing the fraudulent basis of creation science, as a result of which an action group has been formed in resistance to the creationist campaign.

The action group has adopted the name Australian APE (Australian Association for the Protection of Evolution), and aims publicly to refute the claims made by creationists and to press for the removal of creationism from the school science syllabus. It is assembling "survival kits" for schoolteachers who find themselves unwillingly obliged to teach creationism.

Tony Thulborn

Soviet computers

Keeping up with the students

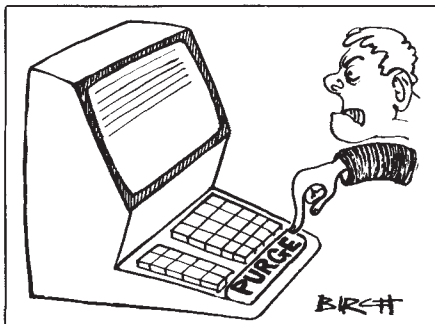
THE Soviet Union has launched a major programme for computer literacy in schools, apparently in response to complaints by Dr Vyacheslav Elyutin, Minister of Higher and Specialized Secondary Education, that the poor performance of certain sectors of the Soviet economy is attributable to the lack of computer-mindedness among middle managers. During the coming academic year, a course in "Fundamentals of information science and computer technology" is to be introduced in all secondary schools, initially in the ninth and tenth forms (for 15 to 16 year olds), and

from another Moscow secondary school using computers and visual display units under supervision and even making microcalculators and assembling printed circuit boards for local industry as part of their "work training experience".

The logistical problems of the new courses are considerable. Soviet school textbooks are usually published in editions of 4 million and software packages will have to be produced on the same scale. Dr Monakhov last month visited the United Kingdom for an international study group on the introduction of computers in schools but, on his return, told a Moscow radio reporter that although he had had considerable opportunity to get acquainted with teaching software of British companies, including Acorn and Sinclair, he found them for the most part unsuitable for Soviet use. The Soviet educationalists were therefore, he said, faced with the "tremendous" task of developing their own software.

Hardware is also liable to cause problems. It is not simply a matter of choosing the computer that is "most suitable" to the teaching profession, Monakhov explained, but also to ensure that Soviet industry is capable of producing such a computer by September. The plans for training of teachers for the new courses also seem unrealistic. Crash courses for teachers of mathematics and physics will be organized (presumably during the summer vacation). Since the Politburo decision also supposes a vast expansion of extra-curricular computer facilities for young people, organized through creativity clubs and the Pioneer youth movement, many teachers may well be faced with pupils more familiar with computers than they are themselves.

Vera Rich



later for younger students. Experiments are also planned in the use of computers in teaching other school subjects, as for example in processing the data obtained in physics or chemistry lessons.

Experiments in the use of computers in schools seem to have been going on for some time. According to Professor Vadim Monakhov, director of the Institute of Teaching Content and Methods of the Soviet Academy of Pedagogic Sciences, one such course was introduced in a Moscow school (specializing in mathematics) as early as 1960. Shortly before the recent Politburo decision, a Soviet television feature showed children

Soldiers' sickness

Compensation for US veterans

Washington

ACTING under instructions from Congress, the US Veterans Administration (VA) has now spelled out its proposals for regulating compensation claims based on exposure to dioxin-contaminated herbicides and radiation from atomic tests. But VA's proposals, which were required to be based on "sound scientific and medical evidence", are unlikely to end the long-running debate over when cause and effect should be inferred in dealing with such claims.

Nineteen million gallons of dioxin-contaminated herbicides were used during the Vietnam war, but veterans have been largely unsuccessful in claiming compensation from the government for disabilities allegedly resulting from their exposure. Only 20 claims have so far been recognized.

VA's proposed regulations continue its policy of assuming some dioxin exposure for all Vietnam veterans, but are unlikely greatly to increase the number of successful claims. Chloracne is the only disease recognized by VA as likely to result from exposure to the most potent dioxin isomer, 2,3,7,8-TCDD, and then only if symptoms appear within 3 months of military service.

Generally, compensation is paid for all medical conditions which appeared during service, and for some chronic diseases appearing within a year of discharge. In borderline cases, compensation is awarded if there is reasonable doubt. Most claims fall outside these categories, however, as they are based on diseases appearing years later.

VA was requested by Congress to consider specifically how it would deal with claims linking dioxin exposure with soft tissue sarcoma and porphyria cutanea tarda, a liver disorder, but has concluded that neither disease has been convincingly shown to result from dioxin.

Most of the evidence on the human health effects of dioxin comes from studies of industrial exposure and from the Air Force's Ranch Hand study, which found some suggestive but equivocal evidence linking relatively high levels of dioxin exposure in spraying crews to increased incidence of skin cancer, birth defects and neonatal deaths.

Firmer answers should emerge from a \$75 million epidemiological study by the Centers for Disease Control (CDC) of health effects of Vietnam service and dioxin exposure, recently mounted. Some 30,000 veterans will be interviewed by telephone, of whom 10,000 will be given thorough physical examinations at special clinical facilities constructed at Albuquerque, New Mexico. CDC are also carrying out a case control study of those particular cancers suspected of having being caused by dioxin. The results should be available before the end of the decade.

Veterans who participated in above-ground nuclear weapons tests in Nevada

and in the Pacific have been just as unsuccessful in prosecuting compensation claims as the Vietnam veterans. Only 20 cancer cases have been recognized as due to service-related exposure to ionizing radiation. VA's proposed regulations allow possible claims based on leukaemias and bone cancer appearing more than two but less than 30 years after exposure (excluding chronic lymphatic leukaemia), and those based on thyroid, breast, lung, liver and skin cancers occurring more than 10 years after exposure. But, because many of these conditions are relatively common in the general population, few cases are likely to be recognized.

The central argument in radiation-linked claims has always been over the size of the estimated dose. Estimates based on film-badge records, for example, take no account of ingested and inhaled radioactivity, which were probably not negligible in tests where soldiers had to run towards a bomb crater throwing hand grenades. The Defense Nuclear Agency has attempted to fill the gap by making estimates of dose

through these pathways, and these are used in a major study of mortality among atomic test veterans that is to be released later this month by the National Research Council's Medical Follow-up Agency. This study, directed by Seymour Jablon, traces mortality among 50,000 test veterans, roughly one-fifth of the total.

But even this will not be enough to satisfy veterans' organizations and the Congress. Acting again under instructions from Capitol Hill, VA is considering the feasibility of an epidemiological survey of test veterans to determine the incidence of morbidity. VA's deliberations are being overseen by Congress's Office of Technology Assessment (OTA), which will shortly publish a feasibility study on the survey. According to Michael Gough of OTA, the Defense Nuclear Agency's dose estimates are likely to be challenged by veterans' representatives, so the question boils down to by how much the dose estimates must be wrong for detectable excess cancers to occur. Whatever OTA recommends, however, the final decision on the epidemiological survey will rest with VA, which may well decide to proceed with the study in order to silence criticism.

Tim Beardsley

Environment

Japanese move to save forests

Tokyo

A GROUP consisting largely of academics has formed in Japan, the world's first society for the promotion of "Green Civilization", which it is hoped will grow into an international organization. Founder members include Kenichi Fukui, Nobel prize-winner in chemistry, and Mostafa Tolba, secretary-general of the United Nations Environment Program.

The organization has been set up largely because of concern over increasing deforestation and desertification of the Earth through poor agricultural practices and industrial pollution. The move is significant in that Japanese companies have been heavily involved in the destruction of tropical rain forest in South-East Asia. Indeed, the nation was dubbed an "ecological menace" in a report adopted last month by the environment commission of the European Communities for "pillaging natural resources in the most irresponsible way possible, and ravaging tropical rain forests for hard woods".

During Japan's "high-growth" phase, huge tracts of land were stripped of all vegetation in many parts of South-East Asia to provide building materials and pulp for Japan; although direct involvement by Japanese companies in the worst of these practices has now lessened, the amount of tropical forest lost each year has increased and Japan remains the major Asian market. The Green Civilization Society says it intends to try to put the ecological point of view over to industrial companies involv-

ed in logging but it is clearly not about to become a grass roots political pressure group.

Unlike Friends of the Earth, much of whose ideology it appears to share, the Green Civilization Society is to be a more academic body. It is believed that by building an interdisciplinary academic movement — for example of ecologists, biotechnologists and forestry researchers — more rational and less destructive means of forest management could be found that could be incorporated into recommendations to the government. Within Japan itself, forestry is regulated by laws devised in the nineteenth century which were meant to provide a balance between exploitation and regeneration of the forest. These laws are no longer effective but will only be changed, Green Civilization members say, by a joint effort between ecologists, forestry researchers and lawyers.

The society has now around a hundred members, most of them drawn from the universities. At the inaugural meeting last week, Seiji Kaya, former president of the Tokyo University, was elected chairman and Michio Tsutsui, professor of forestry at Tokyo University, vice-chairman. So far, funds are being raised mainly from individuals, with some contributions from industry. One hope is that enough money can be raised to provide a "Green Nobel Prize". The establishment of the new group coincides with the designation of this year as International Forestation Year by the United Nations.

Alun Anderson

Japanese tokamak

Plasma now in production

Tokyo

PLASMA was successfully produced in Japan's giant JT-60 tokamak fusion machine for the first time last week. Although this milestone has long since been passed by both of JT-60's competitors — the Tokamak Fusion Test Reactor (TFTR) at Princeton in the United States and the Joint European Torus (JET) at Culham in England — any of the three might be first sometime around 1987 to reach the immediate goal of break-even.

JT-60 scientists were clearly very pleased at how smoothly the tokamak had run. Operation began in April and produced plasma exactly as expected almost immediately. The machine was then run in a discharge cleaning mode to remove impurities from the inner sides of the vacuum chamber and shut down for the installation of diagnostic apparatus which will be used to monitor plasma parameters in the next phase of development.

In the first test run, ohmic heating alone was used and temperatures in the 100-200,000°C range achieved for around 80 milliseconds — a long way from the 100 million degrees and 1-second confinement needed to achieve the break-even condition when the reactor produces as much energy as it consumes. But to achieve that goal, JT-60 will not be following quite the same route as either TFTR or JET.

Unlike either of its competitors, JT-60 will use radio-frequency waves to drive current in the plasma. In the other designs, currents are induced to flow in the plasma by a set of coils around the tokamak that act as the primary winding of a d.c. transformer — the plasma itself acts as the secondary winding. But such a system allows only pulses of current to be produced, for in a d.c. transformer the current in the primary would have to increase indefinitely to continue to drive current in the secondary, a clearly impossible condition. Driving the current by radio-frequency waves, however, allows operation to continue as long as needed as electrons are pushed along the magnetic field by transfer of the waves' momentum.

The JT-60 team is also proud of the neutral beam injector equipment that will be used to heat the plasma in the future. All three tokamaks are using neutral beams but, according to Masaji Yoshikawa, director of Large Tokamak Research at JT-60, the Japanese designs are at least a year ahead of the others in their technical development.

JT-60 also incorporates a unique method of dealing with impurities — metal atoms or gas molecules that may be knocked out of the tokamak inner wall by the plasma — in its magnetic limiter or divertor, and if it works as predicted, will provide a crucial advance towards practical fusion. The idea is to give the plasma in the

tokamak a figure of eight, rather than circular, cross-section, and to keep impurities away from the main body of the plasma in the upper "circle" of the figure of eight.

JT-60 is now expected to be run until the end of June with ohmic heating and then shut down for six months while both neutral beam and radio-frequency heating systems are installed. Next year, power will be put into the plasma from all the systems and, if all goes well, break-even will be reached in the winter of 1987.

All these experiments are to be run with a hydrogen-deuterium plasma; there is no intention to switch to the deuterium-tritium

plasma that would be used in a commercial reactor and will be the subject of second-stage experiments for both TFTR and JET. Neutrons produced in the latter reaction are very energetic and make the whole tokamak radioactive; expensive and difficult remote handling techniques thus become essential.

The Japanese are leaving this stage to their next reactor, the fusion experimental reactor, which will follow JT-60. That is, rather, if some form of massive international effort does not follow on from the three projects now running. Increased costs are making the advantages of international cooperation more and more obvious and recent talks have produced hopes that it might eventually take place.

Alun Anderson

Hungarian education

New law spawns discontent

HUNGARY's new education law, passed last month, lays special stress on the unity of education and research, and the need to provide more flexible forms of post-university training, according to the Minister of Education and Culture, Mr Bela Koepeczi. Presenting the bill to the National Assembly (Parliament), Koepeczi said that although it is not the task of the law to lay down specific conditions for higher and postgraduate education, the new law makes it clear that the provision of proper material conditions in the education sector is the duty of the government. And the government, he assured the assembly, has undertaken to continue its investment in the reconstruction of the universities and to improve the supply of equipment and other necessities for university-based research.

The new law, which replaces the existing legislation of 1961, covers the whole educational process from kindergarten to postgraduate and is intended not as a radical reform but as a gradual process of development, Koepeczi explained. The prospect of new legislation, however, triggered a debate in the universities on academic autonomy, and, in particular, the allocation of senior appointments. A number of academics, Koepeczi said, have gone so far as to press for "decentralization". This term is at present used in Hungarian political circles to mean the reform of 1982, by which industrial enterprises elect their own managers (choosing from at least two candidates). In other words, these scholars advocate that the universities should be allowed to elect their own rectors and senior administrators.

This, Koepeczi says, is not a true "materialization of democracy" since the appointments in question are the concern not only of the institutions individually, but also of the wider scientific and professional community. Although, therefore, the new law gives the universities considerably more control over syllabuses, teaching methods

and day-to-day administration, senior appointments will still be the concern of the government or of the relevant minister. "Democratization" will continue to be ensured by the government or minister consulting with the relevant committees of the Hungarian Academy of Sciences, which thus serve as a kind of watch-dog for scientific public opinion.

Since, in Hungary, members and researchers from the Academy of Sciences frequently simultaneously hold university posts, the system has so far worked reasonably well. The "decentralization" lobby, therefore, seems to have based its stand rather on a matter of principle than on the need to reform any current shortcomings.

This inevitably evokes the question whether the "decentralizers" have not, to some extent, been influenced by the present endeavours of the Polish universities to retain the autonomy won in 1980-1981, in which the right of the universities to elect their own rectors and deans is one of the most burning issues.

Naturally, Mr Koepeczi made no reference to the Polish situation in his address to the National Assembly. The new law was presented as a purely Hungarian concern. Nevertheless, on one point at least, those responsible for drafting the law seem to have had an eye on developments in Poland. In drawing up the bill, the government, Koepeczi said, considered it extremely important to provide a legal background to the operation of student "self-government bodies", an issue that has so far aroused no great controversy in Hungary, but which in Poland is at present a major point of confrontation between the vast majority of the student body (over 90 per cent in some universities) and the governmental Committee for Social and Political Questions which seeks to reduce the self-government committees to a mere adjunct of the Party youth movement.

Vera Rich

The embryo's right to protection

SIR — The opponents of all human embryo research seem to have shifted ground, from the meaningless proposition "life begins at fertilization" to the argument that a human embryo is inherently entitled to full human rights by dint of genetic uniqueness and potential. R. Watson (*Nature* 7 March, p. 10) goes so far as to assert that there can be "no objections" to the idea that a unique potential individual is created "at fertilization".

Leaving aside the difficult questions of when fertilization can be said to have occurred (at sperm penetration, at pronuclear formation, or at activation of the paternal genome some hours or days later?), and of clear differences in potential between embryos, the assertion does not bear close scrutiny. If anything, the beginnings of genetic uniqueness occur in the forming gametes, during the second meiotic division. Here genetic information is exchanged between pairs of homologous chromosomes (the chromosomes of the embryo's grandparents, in fact): in the egg this will have occurred several hours, and in the spermatozoon, several weeks before fertilization. If the embryo is to be granted "rights" on the grounds of genetic uniqueness and potential, how can we logically withhold similar "rights" from the gametes? What of the "rights" of the polar bodies, so wantonly cast aside during oogenesis, and yet clearly of human origin and containing genetic information capable, if suitably nurtured and combined, of potential humanity? One could push the argument for "rights" back still further, to parental embryos, when the germ cell lineage separates from that of the somatic tissues: who can deny that to destroy a primordial germ cell prevents it from fulfilling its life-transmitting potential?

My purpose is not to argue that embryo research should not be constrained, nor that the embryo is worthless. Rather, it is to point out, yet again, that the transmission of life and of human identity forms a continuum, and that there is no point at which human potential suddenly jumps into existence. Setting such arbitrary points may give us the warm glow of moral certainty, but the exercise is not supported by the real complexity of biological systems.

J.M. CUMMINS

*Reproductive Biology Group,
Department of Veterinary Anatomy,
University of Queensland,
St Lucia, Queensland,
Australia 4067*

SIR — The legal and moral status of embryos and the way in which they are handled are the subjects of a complicated debate. R. Watson (*Nature* 7 March, p. 10) seeks to simplify the topic by referring to a clear topological event, fertilization, "at which a unique member of our species is

created". One problem with such an oversimplification is that it may be applied to give silly conclusions.

I contend that each member of a pair of identical twins is a unique member of our species, even though both result from one fertilization. I believe that it is (and should be) illegal to murder one, or to use him or her destructively in medical research. I reach the decision that it is morally wrong to kill a member of a pair of identical twins by deciding that each is a separate, thinking, feeling person, and not by some simple topological argument about the moment when a genome is created.

Of course one might argue that any mitosis (like that which produces identical twins) is an event "at which a unique potential member of our species is created", since it is probably possible to freeze any human diploid cell, await the development of practical human cloning by nucleus transfer, and make an individual of that cell. The killing of single human cells is neither legally nor normally the equivalent of homicide nor, I contend, should it be. This line of argument, too, is an oversimplification.

Whether the benefits of research on embryos (or of abortions) outweigh the disadvantages of destroying embryos is a complex and difficult question. We should address that complex question without looking for quick and easy definitions of what is a person.

JOE WOLFE

*School of Physics,
University of New South Wales,
PO Box 1, Kensington, NSW,
Australia 2033*

Genes and intelligence

SIR — Summarized briefly, John Hartung's hypothesis (*Nature* 311, 515; 1984) is that ill health caused by genetic factors may indirectly affect the intelligence. The principle point of my letter (*Nature* 313, 425; 1985) was to question that *a priori* assumptions that intelligence and school attendance are linked and that they have fewer impair intellectual performance. Both these assumptions may be true, but they are unproven. Since Dr Hartung's hypothesis depends directly on the validity of these assumptions and since they may be tested empirically, it is essential to his argument that they should be examined experimentally.

His reply (*Nature* 314, 398; 1985) deals largely with the viscosity of gases. Of course he is right that the viscosity of gases increases with temperature; he was wrong in his original letter to suggest that humidity also increases viscosity. The problem is, however, far more complex. For instance, the air flow in the nose is turbulent and resistance will therefore depend on gas density, being in general

terms less in humid air than in dry air. The shape of the nose is obviously not wholly dependent on climate. This is illustrated by the observation that the inhabitants of the dry desert regions of North Africa and the Middle East, where natural conditions provide the most viscous air to be found on Earth, are not noted for short noses with wide nostrils.

I agree with Dr Hartung that the search for genetic differences that may also affect intelligence is a proper subject for scientific enquiry. My familial credentials are even more extensive than his, since I can claim to be the son and nephew of both men and women who served in the fight against Hitler's Germany. I think, though, that Dr Hartung would agree with me that neither moral nor political rectitude are genetically determined characteristics. The assertion of familial virtue is no guarantee of the value of his opinions or of mine.

P. M. GAYLARDE

*Royal Free Hospital,
Pond Street,
London NW3 2QG, UK*

Goldfarb's letter

SIR — In your latest article on the Goldfarb case (*Nature* 28 March, p. 307), my name is mentioned as follows: "During his recent conversation with KGB officer Gusev, Dr Goldfarb was confronted with a copy of a letter he had written to Professor Elie Wollman of the Institut Pasteur in Paris, supporting the idea of the moratorium." Gusev said that the letter might constitute a case of 'anti-Soviet propaganda' and made Dr Goldfarb sign a document to that effect".

I wish to make a few points clear.

- I have known David Goldfarb for many years, because he was a pioneer in introducing bacteriophage and bacterial genetics research in the Soviet Union. I consider myself a close friend of this exceptional man, both professionally and personally.

- Our correspondence is brief, containing only personal news of the usual kind, and neither him nor I ever approached in writing any subject that could be interpreted as being controversial in the broad sense of the word.

- I have received no letter from Goldfarb for about a year and was astonished at not even getting a Christmas card, although I wrote to him at the New Year.

- I never received the letter mentioned in your article. There are therefore three possibilities:

- (a) such a letter does not exist and has never existed;

- (b) the letter exists and has been diverted by the Soviet secret police;

- (c) a fake letter has been manufactured.

Hypothesis (b) seems to me to be excluded for the reasons given above.

ELIE WOLLMAN

*Institute Pasteur,
28 Rue du Dr Roux,
75724 Paris Cedex 15, France*

Black hole at the galactic centre

The latest accurate measurements of the size of the galactic core at centimetre radio-wavelengths leave no reasonable escape from the conclusion that there is a black hole at the centre.

WE HAD best learn to live with the idea that there is a black hole at the centre of our Galaxy. That is the implication of the article by K.Y. Lo *et al.* on p.124 of this issue. Perceptive readers will think the paper, which is the culmination of eight years of observation of the galactic centre with coordinated radiotelescopes separated by great distances, is over-modest. Others may think it over-cautious. The phrase "black hole" appears only once, and then as a quotation from an article by Lynden-Bell and Rees, who first floated the idea.

The new development is essentially nothing but a more accurate measurement of the diameter of the radio source known, for the past decade and more, to be located at the galactic centre. The result is startling. At the shorter of the two radio-wavelengths (1.35 cm) at which measurements have been made, the angular diameter of the source is merely 0.002 second of arc (2.1 milli-arc seconds), which at the distance of the galactic centre corresponds to a physical diameter of about 20 astronomical units, the size of the Solar System within the orbit of Saturn. At the longer wavelength (3.6 cm), the source at the centre of the Galaxy appears to be shaped like an ellipsoid with a long axis lying only 8 degrees away from the rotational axis of the Galaxy as a whole.

The case for supposing that there is a black hole at the centre rests almost entirely on the difficulty of accounting for such a prolific source of radio emission in any other way. Lo *et al.* argue that the supposition that the galactic centre might be the site of a rapid burst of star formation, tenable when the limits on the size of the central source were more than a hundred times greater than they have now become, plainly cannot be accommodated within a region no bigger than the orbit of Saturn. A stellar system which is a powerful source of radio emission might fill the bill, but only artificially; such systems are short-lived compared with the lifetime of the Galaxy. Black holes, on the other hand, endure (unless they are very small compared with the masses of typical stars) and even grow more massive with the passage of time. By default, the presence of a black hole at the galactic centre emerges as the simplest explanation of the new observation.

But is it not disconcerting that such an important and interesting conclusion should rest on such a negative argument? That is a natural but not a particularly pointed question. Black holes as such are by their nature invisible, although that does

not apply to whatever matter may be dragged by gravitation into such an entity. Captured material, whatever its chemical composition, will be unspecifically ionized and will yield synchrotron radiation if there are magnetic fields to play with. Visible radiation will however be absorbed (or converted into low-frequency radiation) by the molecular and dust clouds lying around the galactic centre. Infrared observations are also limited in this way, as well as by the mismatch between the resolution of the telescopes for the time being available and the angular diameter of the source. This is why long-baseline radio-interferometry is, for the time being, the best hope for an unambiguous description of the object at the galactic centre.

In the circumstances, the temperature of the source estimated by Lo *et al.* means very little, for the radiation distribution is clearly very different from that of a black body. But the temperature (100 million K) is an accurate measure of the intensity of the radio-emission from the central object and its surroundings. The variation of the apparent angular size of the object with the wavelength of observation is surprisingly tidy — a simple proportionality to the square of the wavelength. The interaction of radiation with electrons is implicated in some way, but just what form this takes will depend, Lo *et al.* say, on the construction of a brightness map for the source.

There is little that can at present be said about the mass of the black hole at the centre of the Galaxy, and indeed a mere measurement of the geometrical size of the region from which radiation is emitted can be neither here nor there. But measurements such as have been attempted in the past few years of the motion of objects, stars and clouds in the neighbourhood of the centre should eventually help to define the mass of the object. As things are, systematic studies of the velocity distribution of nearby objects are inconclusive and open to conflicting interpretations. But presumably it is a matter of time spent on observations before something can be said about the variation with distance of angular velocities of nearby objects about the centre of the Galaxy, from which it should be possible to infer what mass there is at the centre.

Meanwhile, studies of this kind are for some time likely to be more profitable when applied to other galaxies than this. So much should be clear from the account two weeks ago by C.M. Gaskell (*Nature* 314, 672;

1985) of the recent demonstration by J.L. Tonry that there is a black hole at the centre of the dwarf galaxy M32, a satellite of M31 (Andromeda). The technique is to determine the non-radial velocity of central objects as a function of radial distance from the centre, which should be a measure of the mass lying within that distance. Gaskell explained some of the pitfalls in this process, and in particular the failure by these means to demonstrate the presence of a black hole in M87.

The implications of these developments for the present understanding of galactic evolution are clearly important. Any galaxy, if left alone, should eventually collapse under its own weight after several repetitions of the cycle of stellar nuclear synthesis. So there should be no surprise that the end point of galactic evolution should often be a black hole, one whose mass is not very different from that of the original galaxy. That black holes begin forming (at least in M32 and in this Galaxy) is not surprising either, but it is helpful that at least two cases of this phenomenon are now known. To generalize from two cases, one an elliptical galaxy (M32) and one a spiral (this), may seem rash, but it would be no great upset for the appalcarts if most galaxies turn out to have a black hole at the centre.

As is now generally understood, this seems to be the case for active galaxies, quasars spectacularly but other active galaxies such as Seyfert galaxies as well. But it remains to be told whether the difference between such a galaxy and this, where the central black hole is only a modest source of radiation, lies simply in the amount of material accessible to the central attracting source or whether there is something extra special about quasars and similar galaxies. The importance of answering that question is that it will decide whether quasars are exceptional galaxies rather than ordinary galaxies in a particular (and brief) stage in their evolution.

Meanwhile, this development has no immediate bearing on the grander cosmological question of whether there is enough hidden mass to close the Universe, or eventually bring the present expansion to a halt. The mass of the black hole in M32 is estimated at 5 million solar masses, a tiny fraction of the whole. The black hole at the centre of this Galaxy, whatever its mass, is unlikely to be much bigger. So the missing mass that matters must still be sought elsewhere.

John Maddox

Neurobiology

Why flying locusts do not crash

from Jonathan Bacon

LOCUSTS manage to remain aloft for long periods while avoiding crashing into obstacles, their companions in the swarm, or the ground. Studying how they achieve this has kept some biologists busy for almost 40 years but only comparatively recently has the technology been available to examine the neuronal interactions that control these manoeuvres. Intracellular recordings can now be obtained from restrained locusts that still produce the motor rhythm of flight. A paper on page 142 of this issue¹ describes how sensory interneurons that convey information about flight-course deviation interact with the neurones of the flight motor to mediate steering responses.

When a locust is attached to a support so that the legs do not contact the ground, it will 'fly' for as long as a windstream is blown on its head². The early observations were made in the Sherringtonian climate of the time, when it was thought that sensory input was needed to stimulate each reflex burst of motor activity. In terms of the locust flight motor, this implies that each beat of the wing would produce the necessary feedback to stimulate the next wing beat. In 1961, D.M. Wilson performed one of the classical experiments that was to change this way of thinking³. By systematically removing the wings, the top half of the thorax, the abdomen and even the head, he showed that what remained of the nervous system (the thoracic nerve centres, or ganglia that contain the flight motor neurones) was still able to produce a burst of motor activity in response to a puff of wind or an electric shock, although no rhythmic sensory feedback remained. This demonstrated irrefutably that the animal can produce a patterned motor output in the absence of patterned sensory input and established the locust flight motor as a classical example of a central pattern generator (CPG).

Although sensory inputs are patently not required for the locust flight CPG to function in the laboratory, they are necessary in proper flight where turbulence is encountered and the ground and other obstacles are to be avoided. Two sensory cues that help the animal to fly a straight, level course are wind direction and the position of the horizon. Wind direction is monitored by banks of tiny hairs on the head, which the animal uses to determine the attitude of its flying body with respect to the air mass. By constantly making tiny adjustments to its direction of flight through changes in wing stroke, the animal strives to maintain a symmetrical air flow over its head^{2,4}.

Apparent changes in the slope of the horizon are detected by the large compound

eyes and the ocelli, simple light-detecting organs on the head. If the horizon suddenly moves upwards in the animal's field of view, the wing beats are modified to correct the dive. Similarly, L.J. Goodman reported that when the horizon appears to slope to the right, the wing beats are adjusted until the horizon again appears horizontal⁵. The horizon stimulus is very powerful — a locust flying in a device that allows it to rotate freely about its longitudinal axis can be made to fly upside down by inverting the artificial horizon⁵.

Clearly, if the neurones of the thoracic ganglia responsible for producing the correctional changes in the motor output to the wings are to react swiftly to deviations in the flight path, they must be constantly appraised of, among other things, wind direction and horizon position. Several descending sensory interneurons that run from the head ganglion to the thoracic ganglia are candidates for the fast sensory pathways^{6,7}. Early attempts to investigate the cellular basis of steering behaviour sought an explanation in terms of direct connections between these descending interneurons and flight motor neurones, but the synaptic potentials recorded in the motor neurones are very small (and variable) in relation to the massive membrane potential oscillations seen in flight motor neurones in real flight⁸. Although direct synaptic connections between descending interneurons and flight motor neurones certainly exist, they are probably too puny to mediate the rapid steering responses.

Progress in analysing the steering reactions was hampered because the flight motor neurones were the only components of the neural network in the thoracic ganglia that could be identified with certainty. Because sensory neurones are easy to stimulate and muscles easy to record, most workers concentrated on the sensory influences on the flight motor rather than on the working of the motor itself. This contrasts with work on locomotor patterns in soft-bodied invertebrates, such as the leech and the mollusc, where the sensory apparatus is difficult to stimulate precisely and muscles are more difficult to record but the central neurones of the CPG are easy to see and impale.

Further progress could not be made unless the rhythmically active elements of the flight CPG, the cells that must have been active in Wilson's original deafferented preparation³, could be monitored directly. The breakthrough came when R.M. Robertson and K.G. Pearson developed a physiological preparation in which intracellular recordings from thoracic interneurons could be obtained during a burst of activity in the flight CPG⁹. With

much painstaking probing of the neuropile, they have identified sets of interneurons that show rhythmic activity at different phases of the flight cycle, of which some are premotor and some receive sensory inputs¹⁰. A third type belongs to the flight CPG — current injected into such a cell will reset the whole flight rhythm¹⁰. Though much work remains to be done, it seems that a description of the neuronal interactions producing the locust flight CPG is likely to emerge from studying the synaptic interactions between some of these thoracic interneurons¹¹.

Reichert *et al.* have now modified this preparation for investigating the neuronal control of flight steering¹. They provide the locust with a combined horizon and wind stimulus that can be moved to simulate the pitches, rolls and yaws that the animal would experience in flight. The authors reliably identify three pairs of descending interneurons that convey sensory information from the head to the flight circuits of the thoracic ganglia. Each neurone responds to different combinations of wind and visual stimuli with an output that has specific directional characteristics, termed by the authors the "emergent modality" of the neurone.

Even more interesting is the interaction of the output from these deviation detector neurones with thoracic flight neurones. Essentially, Reichert and his colleagues have repeated Goodman's experiment but with intracellular monitoring of interneuronal activity. If the animal could break free of the pins holding it stable for intracellular recording, the changes in flight motor output would enable it to correct for the movement of the horizon stimulus.

The descending neurones make only direct connections with flight interneurons but these have strong inputs to a rhythmically active set of thoracic interneurons. This set also receives a large input from the flight CPG and is presynaptic to flight motor neurones. They therefore effectively act to gate the tonic output of the descending sensory interneurons, so that it is transformed into a phasic input to the flight motor.

Although this newly described mechanism is important because it ensures that the signals from the deviation detectors are delivered to the relevant motor neurones only in their active phase, it is not certain that the information conveyed by the deviation detectors is "intrinsically phase in-

1. Reichert, H., Rowell, C.H.F. & Griss, C. *Nature* **315**, 142 (1985).

2. Weis-Fogh, T. *Nature* **164**, 873 (1949).

3. Wilson, D.M. *J. exp. Biol.* **38**, 471 (1961).

4. Zarnack, W. & Möhl, B. *J. comp. Physiol.* **118**, 215 (1977).

5. Goodman, L.J. *J. exp. Biol.* **42**, 385 (1965).

6. Bacon, J. & Tyrer, N.M. *J. comp. Physiol.* **126**, 317 (1978).

7. Simmons, P.J. *J. exp. Biol.* **85**, 281 (1980).

8. Hedwig, B. & Pearson, K.G. *J. comp. Physiol.* **154**, 745 (1984).

9. Robertson, R.M. & Pearson, K.G. *J. comp. Physiol.* **146**, 311 (1982).

10. Robertson, R.M. & Pearson, K.G. *J. comp. Neurol.* **215**, 33 (1983).

11. Robertson, R.M. & Pearson, K.G. *J. Neurophysiol.* **53**, 110 (1985).

12. Bacon, J. & Möhl, B. *J. comp. Physiol.* **150**, 439 (1983).

dependent" in real flight. At least one pair of wind-sensitive interneurons in the locust is known to be phasically active in flight because the animal's own wing movements inevitably produce phasic wind turbulences around the head¹². Since the three deviation detectors are also wind sensitive, they too may well be modulated at the flight frequency and so the information they carry about horizon movement would already be phasically modulated. However, this modulation would be lost in a fixed preparation.

While work on this kind of preparation is the only way to make progress in understanding flight stabilization behaviour, it does suffer the limitations inherent in im-

mobilizing an animal to study locomotor behaviour. Inevitably, the feedback from the environment is very reduced or completely missing. Any study in neuroethology is bound to be an exercise in compromise because intracellular recordings cannot be made in freely behaving animals and intracellular preparations cannot freely behave. Nevertheless, this work on restrained flying preparations is an important step in understanding the neural circuitry that converts sensory signals to motor output. □

Jonathan Bacon is in the Genetics and Development Group in the School of Biological Sciences, University of Sussex, Brighton BN1 9QN, UK.

Superconductors

Induction of superconductivity by applied magnetic fields

from M. Brian Maple

PHYSICISTS have been fascinated with the interplay between superconductivity and magnetism for nearly three decades, during which period many new and interesting phenomena have been reported (see ref. 1). One of the latest, reported by H.W. Meul and colleagues², is the induction by a high magnetic field of superconductivity in europium-containing metallic compounds of the type $\text{Eu}_x\text{Sn}_{1-x}\text{Mo}_6\text{S}_8$, where x ranges from 0.7 to 0.8. This is a remarkable observation in view of the fact that applied magnetic fields generally destroy superconductivity. It is perhaps equally remarkable that the mechanism apparently inducing the superconductivity, and known as the exchange field compensation effect, was originally suggested more than twenty years ago by V. Jaccarino and M. Peter³.

Despite previous indications of S-N-S-N (S, superconducting; N, normal) behaviour in related compounds in high magnetic fields^{4,5}, it has not been possible to establish the occurrence of magnetic field-induced superconductivity (MFIS) because of the rather broad transitions encountered. Meul *et al.* overcame this problem when they found that the addition of small amounts of Br or Se to the compounds sharpens the transitions. At the same time, they adjusted the superconducting transition temperature, T_c , to within the narrow range of values required to observe MFIS, which was measured in terms of the electrical resistance, R , which has a finite value in the normal state and vanishes in the superconducting state. The inset in Fig. 1 shows normalized electrical resistance (R/R_N) versus applied magnetic field (H) at several temperatures for one of these compounds — $\text{Eu}_{0.75}\text{Sn}_{0.25}\text{Mo}_6\text{S}_{7.2}\text{Se}_{0.8}$. The data at the lowest temperature (0.37 K) clearly reveal a succession of transitions with increasing magnetic field — from superconducting to normal, to supercon-

ducting again, and finally, back to normal (S-N-S-N). From a series of measurements of R/R_N versus H at various temperatures, a phase diagram in the H - T plane was constructed (Fig. 1); two separate domains of superconductivity are found, one at low fields and another at high fields, whereas in a conventional superconductor only one domain of superconductivity at low fields is present.

In order to understand the sequence of superconducting to normal transitions observed, the nature of the superconducting material must be considered. The $\text{Eu}_x\text{Sn}_{1-x}\text{Mo}_6\text{S}_8$ compounds are type II super-

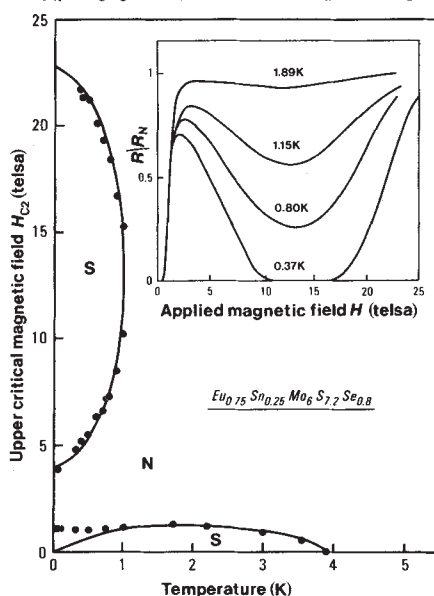


Fig. 1 H - T phase diagram (after ref. 2). The symbols S and N denote superconducting and normal regions, respectively. The solid circles represent the H_{c2} measurements of Meul *et al.*², the solid lines are the boundaries predicted by theory. Normalized electrical resistance, R/R_N , versus applied magnetic field data at several temperatures are shown in the inset.

conductors and hence an applied magnetic field can penetrate into their interior provided it is between the lower critical field value H_{c1} and the upper critical field value H_{c2} , above which superconductivity is destroyed. Under normal circumstances, the value of H_{c2} would be comparable with the paramagnetic limiting field value H_p . However, the $\text{Eu}_x\text{Sn}_{1-x}\text{Mo}_6\text{S}_8$ compounds have magnetic moments associated with the divalent europium ions. Coupling between these magnetic moments and the conduction electron spins generates an 'effective' magnetic field, called the exchange field H_J , that acts on the conduction electron spins in the same manner as an applied magnetic field.

If the sign of the coupling between the Eu magnetic moments and the conduction electron spins is negative, the direction of H will be opposite to that of H_J , that is, the effect of H_J will be 'compensated' by H (see inset to Fig. 2). The net magnetic field H_T is then given by $H_T = H - |H_J|$. The sequence of S-N-S-N transitions shown in the inset to Fig. 1 can be explained in terms of the relationship between H_T and H_p as illustrated schematically in Fig. 2. The upper and lower curves show the applied field H and the exchange field H_J respectively. Summation of these yields the net magnetic field H_T , and it is the variation in this with respect to the value of H_p that determines the superconducting behaviour of the material.

In zero applied field, the compound is superconducting. As the applied magnetic field is increased, the Eu magnetic moments become aligned in the direction of the field at a relatively low H -value and $|H_J|$ attains its maximum value $|H_J^{\max}|$ of 30-50 tesla. Consequently, provided $|H_J|$ is greater than H_p , the magnitude of H_T (which is now negative) can become larger than H_p and drive the material into the normal state. As the applied magnetic field is increased further, the net magnetic field starts to decrease in magnitude, until its magnitude falls below H_p when the compound once more becomes superconducting. Thereafter, H_T increases linearly with H , vanishing and then changing sign at $H = |H_J^{\max}|$ where H_J is completely compensated, until H_T (now positive) exceeds H_p , when the system once again becomes normal. Because the value of the exchange field can be rather large, the compensation effect can take place at very high magnetic fields.

In actuality, the second transition from superconductivity to normal state will also occur if H surpasses the orbital critical field H_{c2}^* of the compound. Thus, in order to observe MFIS, it is important that H_{c2}^* , which can also limit the value of the upper critical field H_{c2} , is large enough to allow superconductivity to occur in this magnetic field range (that is, $H_{c2}^* > |H_J^{\max}| - H_p$), since the orbital effects are not compensated by the Jaccarino-Peter mechanism. The compensation of the exchange field by the applied magnetic field in several

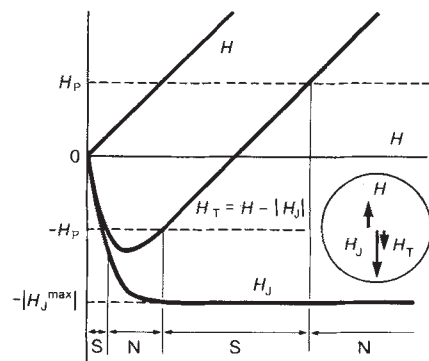


Fig. 2 Schematic diagram of the S-N-S-N transition sequence in an applied magnetic field H . H_T is the net magnetic field acting on the conduction electron spins, H_J is the exchange field produced by the Eu magnetic moments, and H_p is the paramagnetic limiting field. The material is superconducting (S) when $|H_T| < H_p$ and normal (N) when $|H_T| > H_p$.

$\text{Eu}_x\text{M}_{1-x}\text{Mo}_6\text{S}_8$ systems (where M is a metal such as tin, lead, lanthanum⁷ or ytterbium⁸) and in EuMo_6S_8 under pressure⁹, had been inferred from an enhancement and anomalous temperature dependence of H_{c2} in these materials. Independent evidence for a negative coupling between the Eu magnetic moments and the conduction electron spins has emerged from the nuclear magnetic resonance and Mössbauer studies of Fradin *et al.*¹⁰, electron paramagnetic resonance measurements by Odermatt¹¹ and the band structure calculations of Freeman and Jarlborg¹². The phenomenon of MFIS was thus anticipated, but samples of the necessary high quality, in which the candidates for MFIS were satisfied, have not been available until the present study by Meul *et al.*².

Following the initial suggestion by Jaccarino and Peter that MFIS could occur in a ferromagnet due to exchange field compensation, several theoretical investigations of MFIS in paramagnetic systems were carried out¹³⁻¹⁵. One model, based on the theory of type II superconductivity including the effect of the exchange field¹⁴, was used by Meul *et al.* to analyse their $H_{c2}(T)$ measurements on the $\text{Eu}_x\text{Sn}_{1-x}\text{Mo}_6\text{S}_8$

system. The calculated boundaries of the two superconducting domains observed for the compound $\text{Eu}_{0.75}\text{Sn}_{0.75}\text{Mo}_6\text{S}_{7.2}\text{Se}_{0.8}$ are indicated by the solid lines in Fig. 1 where the parameters of the theory (see ref. 2) have been adjusted to give the best fit to the data. The excellent description of the H - T superconducting phase boundaries shown in Fig. 1 within the context of the Jaccarino-Peter compensation mechanism can be regarded as singularly impressive. However, several important issues remain to be resolved before a complete understanding of this phenomenon can be attained. The behaviour of the superconducting energy gap and the flux-line lattice should be characterized, and the thermodynamic order of the magnetic field-induced tran-

sition needs to be established. It would, of course, be interesting to observe MFIS in other materials; preliminary evidence indicates that it may occur in the $\text{Eu}_x\text{Ho}_{1-x}\text{Mo}_6\text{S}_8$ system¹⁶.

Twenty-two years after its prediction by Jaccarino and Peter, MFIS due to the exchange field compensation effect seems to have been verified by Meul *et al.* on a paramagnetic system. But Jaccarino and Peter originally suggested MFIS in a weakly ferromagnetic material, assuming that it would be superconducting in the absence of ferromagnetic ordering. MFIS in a ferromagnet remains to be discovered. □

M. Brian Maple is in the Physics Department, University of California, San Diego, La Jolla, California 92093, USA.

Hybrid oncogenes

A multi-purpose tool for studying multi-step processes

from Peter W. J. Rigby

INTRODUCTION of exogenous DNA into the mouse germ line is a well-established technique that has been applied to many biological problems. Most previous work has concentrated on the analysis of DNA sequences that mediate tissue- and cell-type-specific gene expression¹⁻⁷, but more complex biological phenomena can also be studied using this method^{8,9}. Insertion of oncogenes, of either viral or cellular origin, into the germ line provides an *in vivo* system in which the mechanism of tumorigenesis can be analysed and is also likely to facilitate the study of development, differentiation and cellular interaction. The paper by D. Hanahan on page 115 of this issue¹⁰ describes the most interesting application of this approach so far reported.

Hanahan has constructed hybrid oncogenes in which the coding sequences of simian virus 40 (SV40) large-T antigen — a nuclear DNA-binding protein that can transform various cell types, including primary cells¹¹ — is placed under the control of DNA from the upstream region of the rat insulin gene¹⁰. Transfection experiments using cultured cells have shown that this region contains a cell-type-specific transcriptional control element that functions in insulinoma cells but not in fibroblasts¹². Hanahan has used two types of construct: in one, the insulin upstream sequences are oriented so that the insulin promoter drives transcription of the T-antigen gene; in the other, the upstream sequences have the reverse orientation and so transcription of the viral oncogene cannot be driven by the conventional insulin promoter.

Lines of mice carrying each of these constructs in their germ line have been established. The mice develop normally but die prematurely at around nine to twelve weeks

of age; before death their behaviour is normal. Immunological studies show that expression of large-T antigen is confined to the β -cells of the endocrine pancreas, indicating that the upstream sequences of the insulin gene are sufficient to confer cell-type-specific expression *in vivo* as well as *in vitro*. Proper expression occurs in mice carrying either construct; in the case of the inverted control region it is not known where the messenger RNA initiates.

Histopathological and immunocytochemical analyses show that in young mice the islets of Langerhans are of normal size but that all of the β -cells express large-T antigen whereas the α - and δ -cells, which synthesize glucagon and somatostatin, respectively, are reduced in number and are disordered. With increasing age most of the islets become hyperplastic and are comprised almost entirely of β -cells expressing large-T antigen. Finally, a small percentage of the hyperplastic islets give rise to well-vascularized β -cell tumours which do not metastasize.

Two other groups have recently developed lines of transgenic mice carrying oncogenes. In one study, the oncogene was again that encoding SV40 large-T antigen, but in this case expression was controlled by the virus's own promoter/enhancer region¹³. The mice develop normally but again premature death occurs, mainly as the result of tumours in a particular region of the brain, the choroid plexus. The other group constructed a hybrid oncogene in which murine *c-myc* encoding sequences were placed under the control of the promoter/enhancer in the long terminal repeat of the hormone-responsive mouse mammary tumour virus¹⁴. They observed expression of the hybrid oncogene in various tissues including the mammary gland and, most commonly, the salivary gland. In two

1. Superconductivity in Ternary Compounds, Topics in Current Physics 34 (eds Maple, M.B. & Fischer, Ø.) (Springer, Berlin, 1982).
2. Meul, H.W., Rossel, C., Decroux, M., Fischer, Ø. & Remenyi, G. & Briggs, A. *Phys. Rev. Lett.* **53**, 497 (1984).
3. Jaccarino, V. & Peter, M. *Phys. Rev. Lett.* **9**, 290 (1962).
4. Isino, M., Kobayashi, N. & Muto, Y. in *Ternary Superconductors* (eds Shenoy, G.K., Dunlap, B.D. & Fradin, F.Y.) 95 (North-Holland, Amsterdam, 1981).
5. Wolf, S.A. *et al.* *Phys. Rev.* **B25**, 1990 (1982); Fuller, W. W. *et al.* *J. appl. Phys.* **53**, 2622 (1982).
6. Fischer, Ø. *et al.* *J. Phys.* **C8**, L474 (1975).
7. Torikachvili, M.S. & Maple, M.B. *Solid State Commun.* **40**, 1 (1981).
8. Torikachvili, M.S., Beille, J., Lambert, S.E. & Maple, M.B. in *Superconductivity in d- and f-band Metals* (eds Buckel, W. & Weber, W.) 241 (KFK GmbH, Karlsruhe, 1982).
9. Decroux, M. *et al.* *Phys. Rev. Lett.* **52**, 1563 (1984).
10. Fradin, F.Y. *et al.* *Phys. Rev. Lett.* **38**, 719 (1977).
11. Odermatt, R. *Helv. Phys. Acta* **54**, 1 (1981).
12. Freeman, A.J. & Jarlborg, T. in *Superconductivity in Ternary Compounds, Topics in Current Physics* Vol. 34 (eds Maple, M.B. & Fischer, Ø.) 167 (Springer, Berlin, 1982).
13. Schwartz, B.B. & Gruenberg, L.W. *Phys. Rev.* **177**, 747 (1969).
14. Fischer, Ø. *Helv. Phys. Acta* **45**, 229 (1972).
15. Maekawa, S. & Tachiki, M. *Phys. Rev.* **B18**, 4688 (1978).
16. Capone, D.W. *et al.* *Proc. 30th. A. Conf. Magnetism and Magnetic Materials* (in the press).

lines of mice, adenocarcinomas of mammary origin appeared in multiparous females at the time of pregnancy. The hormonal induction of expression that occurs in the lactating mammary gland is presumably prerequisite for the accumulation of sufficient levels of the *c-myc* oncogene product to induce the first step of tumorigenesis.

Although the *in vitro* experiments that suggest the existence of at least two classes of oncogene required for the transformation of primary cells¹¹⁻²⁰ have aroused much interest, it is important to extend these findings to *in vivo* studies. Hanahan's results show that SV40 large-T antigen induces proliferation of the β -cells and the subsequent hyperplasia but suggest strongly that this is not sufficient to induce a tumour, otherwise all the hyperplastic islets would progress to tumours¹⁰. Similarly, tumours arise in only a few of the mammary glands of the mice carrying mammary tumour virus/*c-myc* oncogene. In these situations some second event must be required for tumorigenesis. An *in vivo* system in which the multi-step nature of tumorigenesis can be analysed has thus been established.

In vitro experiments are severely limited by the availability of suitable lines of cultured cells and by transfection efficiencies, whereas *in vivo* experiments allow expression of the oncogene to be targeted to any cell type for which an appropriate transcriptional control sequence is available. Large-T antigen could thus be expressed in many different cell types to ask whether the second event always affects the same gene or whether there is cell-type specificity for the subsequent event(s). Conversely, the insulin upstream region could be used to target the expression of other oncogenes to the β -cells to ask, first, whether hyperplasia and subsequent tumorigenesis are induced and, second, whether the subsequent event(s) is general, or specific to the initiating oncogene. Finally, two or more oncogenes could be targeted simultaneously to the same cell type. Such studies will greatly advance our understanding of the multi-step nature of oncogenesis. Equally important, they will enable molecular studies on tumour types for which there is no good *in vitro* model. The system should also make possible the study of processes, such as angiogenesis, metastasis and immune surveillance, that are hard to study *in vitro*.

It is curious that the β -cell tumours seem not to induce an immune response. SV40 itself is an extremely weak carcinogen in mice; tumours arise very rarely and with a long latent period^{21,22}, presumably because the small proportion of the large-T antigen that is displayed on the surface elicits a powerful immune response.

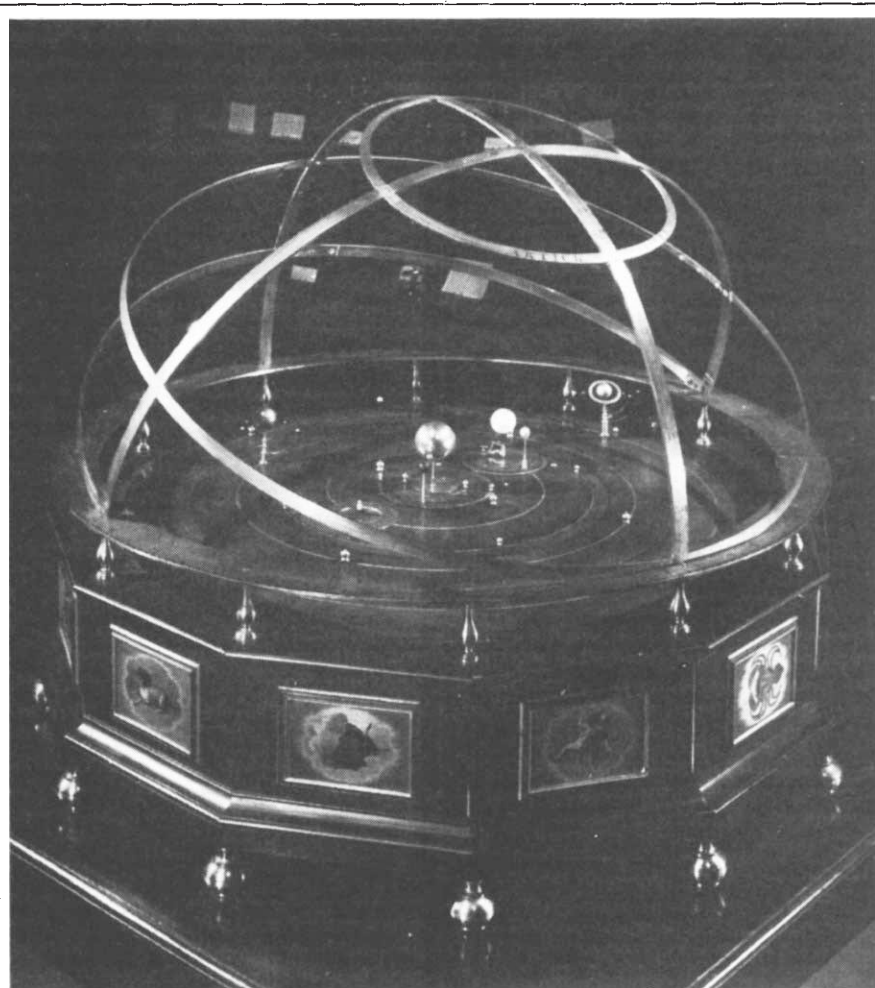
The severe disorganization of the other cell types in the endocrine pancreas induced by β -cell proliferation and hyperplasia is also of interest. The ability to perturb the growth of one cell type in complex

interacting system will allow studies of the mechanisms that normally control the orderly growth and behaviour of an organ.

One further use of this type of study derives from the ability of many oncogenes, including those discussed here, to 'immortalize' cells. Thus, it may be possible to establish cell lines derived from all tissues of mice carrying such genes in their germ line. Indeed, it has already been shown that cells can be readily grown from a number of the tissues of the SV40 transgenic mice (A.J. Levine, personal communication). Such lines have considerable practical use. Commercial manufacture of some proteins may be limited by the inability of commonly used cell types to perform some post-translational modifications. The problem would be overcome if established lines of cells normally making the protein were available. □

1. Brinster, R. *et al.* *Nature* **306**, 332 (1983).
2. Chada, K. *et al.* *Nature*, **314**, 377 (1985).
3. Grosschedl, R. *et al.* *Cell* **38**, 647 (1984).
4. McKnight, G. *et al.* *Cell* **34**, 335 (1983).
5. Shani, M. *Nature* **314**, 283 (1985).
6. Storb, V. *Nature* **310**, 238 (1984).
7. Swift, G. *et al.* *Cell* **38**, 639 (1984).
8. Richie, K.A. *et al.* *Nature* **312**, 517 (1984).
9. Rusconi, S. & Kohler, G. *Nature* **314**, 330 (1984).
10. Hanahan, D. *Nature* **315**, 115 (1985).
11. Rigby, P.W.J. & Lane, D.P. *Adv. viral Oncol.* **3**, 31 (1983).
12. Walker, M. *et al.* *Nature* **306**, 557 (1983).
13. Brinster, R. *et al.* *Cell* **37**, 367 (1984).
14. Stewart, T.A. *et al.* *Cell* **38**, 627 (1984).
15. Fasano, O. *et al.* *Molec. cell. Biol.* **4**, 1695 (1984).
16. Eliyahu, D. *et al.* *Nature* **312**, 312, 646 (1984).
17. Jenkins, J.R. *et al.* *Nature* **312**, 651 (1984).
18. Land, H. *et al.* *Nature* **304**, 596 (1983).
19. Parada, L.F. *et al.* *Nature* **312**, 649 (1984).
20. Ruley, H.E. *Nature* **312**, 602 (1983).
21. Abramczuk, J. *et al.* *J. Virol.* **49**, 540 (1984).
22. Lewis, A.M. & Look, J.L. *Science* **227**, 15 (1985).

Peter W.J. Rigby is in the Cancer Research Campaign's Eukaryotic Molecular Genetics Research Group, Department of Biochemistry, Imperial College of Science and Technology, London SW7 2AZ, UK.



Celestial mechanics

THIS fine example of an orrery, or mechanical model of the Solar System, is one of the early scientific instruments on display in the exhibition "Science and Profit in Eighteenth Century London" at the Whipple Museum of the History of Science in Cambridge until 6th December 1985. Orreries, used to teach the astronomy of the Solar System, were not made to scale but illustrate the relative positions and motions of the planets and their moons. This one, made by Geo Adams at Tycho Brahe's Head in Fleet Street, London, is calibrated for the Julian calendar, which dates it before 1752. It includes Saturn with its rings and the five moons found between 1655 and 1684, but not Uranus, which was discovered only in 1781. Other instruments in the exhibition include an altazimuth theodolite, made about 1740, a compound microscope improved by tilting the limb for extra comfort, and a reflecting telescope.

Circadian rhythms

Period piece for *Drosophila*

from Kevin Moses and Michael Ashburner

NORMAL male fruit flies sing a courtship song to their prospective mates by vibrating their wings. There is a 20–50 ms interval between pulses of sound—and the length of this interval itself oscillates with a period of 50–60 s. This oscillation is under the control¹ of a genetic locus known as *period*, or *per*, which also plays a part in controlling the rhythm of locomotor activity of the flies. It is one of seven known genetic loci that determine the rhythm of eclosion, that is, emergence from the pupa, in a population of flies. The *per* locus was the first of the seven to be discovered, is still the best characterized and has recently yielded to molecular analysis by two groups, with fascinating, if not altogether consistent, results^{2,3}.

Genetic analysis of circadian rhythms of *Drosophila melanogaster* began with Konopka and Benzer's isolation of mutations affecting the eclosion rhythm⁴. Mutations of *per* can have a variety of effects on rhythm, which can become shorter than the normal 24 h—for example, 19 h in *per^S* flies; or longer—for example, 29 h in *per^L* flies; or arrhythmic—for example, in *per⁰* flies. Similarly, the 50–60 s oscillation in the period between bursts of love song in normal male flies is shortened to 40 s in *per^S* flies, lengthened to 80 s in *per^L* flies and absent in *per⁰* flies.

Thus, mutant alleles of *per* affect both circadian rhythms and much faster rhythmic activities. By transplanting brains between flies that differ in their genotype at the *per* locus, Handler and Konopka have shown that it is the expression of this gene in the brain that determines the fly's phenotype⁵. This suggests that the fly's behaviour is continuously affected by the expression of the *per* gene, a conclusion supported by studies with a temperature-sensitive *per⁰* allele (ref. 6). The nature of the mutations that affect this rhythm is interesting. A complete absence of *per* function (as seen in a homozygous *per⁻* deletion) gives an arrhythmic fly, but a partial loss of function gives a fly with a long cycle. This is the conclusion from the observation that females that are hemizygous for *per⁺* have a 25-h cycle.

For their molecular analysis of *per*, the groups from Rockefeller University² and Brandeis University³ have used different starting points. The former cloned *per* by jumping from the previously cloned *Notch* gene, crossing the break points of a deletion that brought *Notch* just distal to *per*. From this starting point, they have characterized about 90 kilobases (kb) of contiguous DNA and by mapping the breakpoints of other chromosome aberrations to the DNA, have defined the *per* locus to a 25-kb interval. On the other

hand, the group from Brandeis University microdissected polytene chromosomes to isolate DNA from band 3B1.2, which contains *per*, for cloning. Both groups have characterized transcripts in this region and have confirmed the cloning of *per* DNA by transformation.

Within the 25 kb of DNA that must, on genetic criteria, include the *per* locus, several transcripts have been identified. The two groups differ somewhat in their interpretation of the transcription data. Between the limits of two deletions that bracket the *per* gene the Brandeis group have found four transcripts, and the Rockefeller group three. Perhaps the most startling discovery by the Brandeis group is that the amount of the 0.9-kb transcript varies during the day: it is more abundant at 2 pm than 2 am. Furthermore *per⁰* flies have a greatly reduced amount of this transcript. The translocation *T(1;4)JC43* provides an interesting mutation of *per* because it acts as a *per⁰* mutation in eclosion assays but as a *per^L* mutation in locomotor-activity assays. This translocation has a correspondingly complex effect on transcripts from the *per* region, abolishing a 4.5-kb transcript, reducing the level of 1.1-kb transcript and producing two novel transcripts of 11.5 and 0.9 kb from the *per* region, according to the Rockefeller group². The Brandeis group also finds two new transcripts (which they call 10 and 0.5 kb) and a reduction,

but not a complete absence, of the 4.5- and 0.9-kb RNAs. Presumably the 0.9-kb Brandeis RNA and 1.1-kb Rockefeller RNA are the same species.

Unfortunately, the transformation data do not help to define the function of the *per* gene transcripts because the DNA used either did not include sequences that code for the cycling 0.9-kb RNA⁷ or included the equivalent DNA but also the DNA that codes for the 4.5- and 1.0-kb RNAs⁸. The ambiguity results from the fact that all of the transformants remain mutant, with longer rhythms than normal. Whether this is due to a failure to obtain correct tissue-specific expression of the introduced *per* DNA, or whether the DNA lacks sequences required for normal *per* function, or is due to differences in the assay of the *per* phenotype, is unclear.

The fact that the two groups disagree as to the details of the structure and function of *per* should not detract from their achievements nor from the great interest in this gene. It is clear that we are now on the way to a molecular description of a gene whose function is required for both circadian and other rhythmic behaviours. □

1. Kyriacou, C.P. & Hall, J.C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6729 (1980).

2. Bargiello, T.A. & Young, M.W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2142 (1984).

3. Reddy, P. *et al. Cell* **38** 701 (1984).

4. Konopka, R. & Benzer, S. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2112 (1971).

5. Handler, A.M. & Konopka, R. *Nature* **279**, 236 (1979).

6. Orr, D.P.-Y. thesis, California Institute of Technology (1982).

7. Bargiello, T.A., Jackson, F.R. & Young, M.W. *Nature* **312**, 752 (1984).

8. Zehring, W.A. *et al. Cell* **39**, 369 (1984).

Kevin Moses and Michael Ashburner are in the Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK.

Archaeology

Using isotope chemistry to detect prehistoric diets.

from Glynn Ll. Isaac

WHEN archaeologists seek to reconstruct the subsistence systems used by prehistoric hunter-gatherers, they face two particularly intractable problems. The first is to quantify the relative importance of plants versus animals in ancient diets, and the second is that of determining patterns of seasonal movements. Studies on bone composition have the potential to solve the second type of problem, as illustrated by the paper of Judith Sealey and Nikolaas van der Merw on page 138 of this issue¹.

In sites where organic remains are relatively well preserved, it is often possible, by patient study of the animal bones and shells, to rank the relative importance of different animal species as a source of food. Sometimes the same can be done for carbonized plant remains. However, it is seldom possible to gauge the proportions of plants and animals relative to one

another. Similarly, while particular remains can sometimes allow archaeologists to say that a site was occupied during one season, it is often much harder to be sure that the site was *not* occupied in other seasons. So, having found a few sites with complementary seasons of apparent occupation, is it safe to link up these observations and infer a patterned seasonal round of movements by the occupants?

During the 1960s and 1970s, the seasonality issue was dramatized by the writings of a Cambridge school of economic prehistorians inspired by Grahame Clark and led by the late Eric Higgs. In effect, Higgs suggested that the most efficient use of many landscapes would have involved sweeping seasonal movements of inhabitants and livestock. The model was (and is) attractive, but how was its applicability to particular pieces of prehistory

to be tested? Did hunter-gathers maximize their foraging? What effects might the packing of several social groups into a landscape have on the feasibility of seasonal movements? These are questions of interest and importance in archaeological modelling of prehistoric ecology, but in very few cases has it been possible to obtain evidence to test intuitive answers.

Fortunately, archaeologists are optimists who continue to pursue their enquiries Micawber-like, and in the last few years they have been rewarded. It has emerged that to some extent humans are what they eat. Diet affects the composition of bone in various ways and subtle signals, indicative of various aspects of the subsistence regime, may be preserved in prehistoric skeletons. For instance, in any given region the calcium to strontium ratio at the time of death can be used to separate pure herbivores from pure carnivores and to establish a scale of mixed diets^{2,4}. In addition, the carbon-13 composition of bones can serve as a measure of the relative dietary contribution of C₃ and C₄ plants; for example, the spread of maize into the woodlands of North America has been monitored in that way^{5,6}. Nitrogen isotopes, along with carbon isotopes, are providing measures of the relative importance of seafoods and legumes in some ancient diets⁷⁻⁹.

Thus, in recent years, aspects of stable isotope chemistry and human bone biology have become important in studies of prehistoric subsistence systems, such as that reported by Sealy and van der Merwe in this issue¹. Their investigation took as its starting point what must surely rank as one of the most sustained and informative archaeological studies of regional hunter-gatherer subsistence and land-use patterns, carried out over some fifteen years by a group at Cape Town University led by John Parkington¹⁰. It has monitored responses that span a sequence of environmental change starting 20,000 years ago. Amongst other things, Parkington suggests that there existed at some stage a strong pattern of seasonal migration of human groups between the coast and an inland range of mountains. A series of sites with evidence of complementary seasons of occupation make the model plausible and attractive. But is it correct?

Sealy and van der Merwe report an ingenious and painstaking investigation of bone composition, which they use to explore the question of seasonal movement. Extensive measurements of a full range of coastal foods show that they have markedly different carbon-13 abundances from a full range of mountain foods (be they plant or animal). From this, it is possible to predict the isotope content of bones from people on a coastal diet, a mountain diet or the mixed diet that would indicate seasonal migration between the coast and inland mountains. Tested against the predictions, the compositions of bones do not sustain the hypothesis of extensive seasonal migration

of the same people between the two habitats.

Although the results hold out promise that bone-composition studies will sometimes provide otherwise unavailable information about prehistoric ranging patterns and the use of food from different ecozones, the findings in this particular case must be seen as provisional. The samples of skeletons from each time period and each zone are small and need enlarging. Careful alternative modelling of various mixed diets will also need to be considered. Nevertheless, the patterns seem consistent enough to be highly suggestive that, attractive though it may have seemed even then, spending summer in the Alps and winter

on the Riviera was not always feasible in prehistoric times. □

1. Sealy, J. & van der Merwe, N.J. *Nature* **315**, 138 (1985).
2. Schoeninger, M.J. *Museum of Anthropology, Univ. Michigan Tech. Rep.* 9 (1979).
3. Schoeninger, M.J. *Paleoethnology* **7**, 73 (1981).
4. Sillen, A. & Kavanagh, M. *Ybk phys. Anthr.* **25**, 67.
5. van der Merwe, N.J. & Vogel, J.C. *Nature* **276**, 815 (1978).
6. van der Merwe, N.J. *Am. Sci.* **70**, 596 (1982).
7. De Niro, M.J. & Epstein, S. *Geochim. cosmochim. Acta* **45**, 341 (1981).
8. Schoeninger, M., De Niro, M.J. & Tauber, H. *Science* **220**, 1381 (1983).
9. Farnsworth, P., Brady, J.E., De Niro, M. & MacNeish, R.S. *Am. Antiq.* **50**, 102 (1985).
10. Parkington, J.E. in *Advances in World Archaeology* (ed. Wendorf, F. & Close, A.) **89** (Academic, London, 1984).

Glynn L.I. Isaac is in the Department of Anthropology, Peabody Museum, Harvard University, Cambridge, Massachusetts 02138, USA.

Meteorology

What limits front formation?

from K.A. Emanuel

FOR many years, meteorologists and oceanographers were perplexed by the tendency of fluid flows to form 'fronts', or near discontinuities, in such properties as temperature, salinity gradients and velocity. These fronts, which occur in both the atmosphere and oceans, seem to defy the general tendency of turbulent flows to smooth out strong gradients of fluid properties. Then, in 1972, B.J. Hoskins and F.P. Bretherton (*J. atmos. Sci.* **29**, 11; 1972) developed an attractive theory which clearly explained the dynamics of front formation and which predicted the formation of real discontinuities in density gradients and velocity. This presented meteorologists with a new problem: why do actual fronts consist of zones of strong gradients, spread out over horizontal distances on the order of several tens of kilometres, rather than of real discontinuities?

It is natural to suppose that turbulent mixing limits the gradients found in real fronts and, until recently, it was assumed that the mature frontal zone represents a balance between the dynamics of front formation and their dissipation by turbulence. Recently, however, I. Orlanski and B.B. Ross (*J. atmos. Sci.* **41**, 1669; 1984) have suggested that frontal collapse may be at least partially limited by processes that do not involve turbulent mixing. Their conclusions stem from the results of numerical simulations of atmospheric fronts, making use of the full equations that describe the conservation of momentum, heat and mass, rather than the approximate forms used in the Hoskins-Bretherton model.

These approximate equations, which can sometimes be solved analytically and which have proved extremely useful in advancing the conceptual picture of front formation, are based on an assumption of cross-front geostrophic balance — that is, forces arising from pressure gradients across the front are assumed to be balanced by Coriolis accelerations acting on the component of

wind parallel to the front. This approximation can be shown to be valid during most of the time it takes for density gradients to collapse but it eventually breaks down near the time of formation of discontinuities. When this happens, the cross-front pressure gradient force may be balanced by actual particle accelerations across the front as well as by Coriolis accelerations. The numerical simulations carried out by Orlanski and Ross seem to show that these effects tend to limit rate of front formation, at least, and perhaps also the ultimate amplitude of the cross-front gradient of front-parallel velocity.

As Orlanski and Ross point out, an exact interpretation of the numerical simulations is problematical. It is generally necessary to represent some turbulent mixing to keep the time integration stable, and it is not clear what role this diffusion plays in limiting the numerically simulated fronts. Detailed evaluation of the dynamics of the simulated fronts is also hampered by the finite horizontal resolution of numerical models. (In this case, the horizontal distance between grid points is 61.5 km.)

Finally, it is well known that near discontinuities can form in non-rotating fluids in which the balance in the frontal zone involves no Coriolis accelerations; taken together with present results this implies that, while both purely inertial and purely geostrophic fronts are limited by turbulent mixing, fronts involving both inertial and Coriolis accelerations may be limited by a non-diffusive mechanism.

This non-intuitive idea, if true, presents a startlingly different view of the dynamics of mature fronts in the atmosphere and oceans. More work with numerical models and better observations of frontal zones should help assess its validity. □

K.A. Emanuel is at the Center for Meteorology and Physical Oceanography, Cambridge, Massachusetts 02139, USA.

Bacterial immunology

Antigenic variation in gonococcal pili explained

from J.R. Saunders

ANTIGENIC variation of surface structures in pathogens is an important mechanism for avoiding host defence mechanisms. *Neisseria gonorrhoeae*, the causative agent of gonorrhoea, is well known for its ability to exhibit interstrain and intrastrain variation in the type of pilus (fimbria) that it produces. Pili — hair-like proteinaceous structures — are important virulence factors in gonococci because they mediate adhesion of the bacterium to mucosal surfaces during the initial stages of infection. Variability in such exposed surface structures goes some way to explaining the success of this bacterium in causing repeated infections in humans. On page 156 of this issue, Per Hagblom and coworkers have shown that pilin, the protein subunit of gonococcal pili, consists of constant and variable regions with the amino-terminus being conserved and the carboxy-terminus containing more-or-less variable amino-acid sequences¹. They propose that antigenic variation in gonococcal pilins can be explained by chromosomal rearrangements that result from intragenic recombination between a repertoire of cassettes of different pilin genes^{1,2}.

Gonococcal pili are composed of identical pilin subunits of relative molecular mass 18,000–22,000, with apparent relative molecular mass and antigenic differences depending on the protein strain involved. The 53 or so amino acids at the amino-terminus of the polypeptide are conserved in all pili from *N. gonorrhoeae* and constitute a constant region^{1,3}. The remaining 100–110 amino acids that form the carboxy-terminus of the protein are variable and account for the antigenic diversity observed within and between strains. This variable region can be divided into two domains, a semi-variable region that extends between amino-acid residues 54 and 114 and contains many different amino-acid substitutions depending on the pilin examined, and a hypervariable region that extends from residue 115 to the carboxy-terminus and contains deletions as well as substitutions and insertions of 1–4 amino acids in different pilin variants². Immunological and sequence studies suggest that the hypervariable region is itself flanked by short (about 10-residue) regions of conserved amino-acid sequence that centre around the cysteine residues at positions 121 and 151 (refs 1,4). Polyclonal antibodies raised against purified pili are primarily directed at antigenic sites that lie within the disulphide loop formed between these cysteine residues⁴. The constant region seems to contain a much more weakly immunogenic epitope(s) that is

common to all gonococcal pili.

Antisera raised against synthetic peptides corresponding in sequence to residues 41–50 in the constant region and residues 69–84 in the semi-variable region of pili from *N. gonorrhoeae* strain MS11 are effective in inhibiting the binding of heterologous gonococci to human cell lines⁴. In contrast, monoclonal antibodies directed against the variable, type-specific hypervariable domain, but not the conserved domain, are able to block binding of purified whole pili from *N. gonorrhoeae* strain P9 to animal cells⁵. These findings suggest that components of both conserved and variable regions are involved in binding to cell receptors because of the way that the mature pilin polypeptide is folded.

Appropriate synthetic peptides, either singly or in combination, might be useful as components of vaccines directed against gonorrhoeae. Such vaccines would work by eliciting antibodies that would prevent adhesion of the pathogen to the epithelial surface. The immunodominance of the variable part of pilin molecules and the many immunologically distinct pili that can be assembled by just one strain of the gonococcus may compromise such an approach.

Variability in piliation of *N. gonorrhoeae* involves two distinct genetic phenomena. In the first, pilus phase variation, pilin genes are reversibly switched on (to produce pilated or P⁺ cells) or off (to produce non-piliated or P[−] cells) at a high rate. The second process, antigenic variation, involves the generation of biochemical and antigenic diversity in pilus type and is caused by the expression of alternative pilin genes. The two processes are probably closely linked as the transition between P⁺ and P[−] states is accompanied by genome rearrangements; P[−] derivatives that switch back to the P⁺ state frequently express pili that are antigenically distinct from those produced by the original P⁺ strain^{1,3}.

The ability to alter piliation status by a phase change is of great advantage to this pathogen because, although they are important adhesions for initiating infection, the pili might render the organism more susceptible to host defence mechanisms once access to tissues has been gained. It is also not yet clear whether the genetic mechanism for antigenic variation requires phase variation to the P[−] state as an obligatory intermediate stage. Pili are major surface antigens in the gonococcus and, therefore, frequent changes in their antigenic nature may assist in avoiding host immune responses. Variations in pili may also be associated with differential attach-

ment to, and virulence for, animal cells, suggesting that one role of antigenic variation is to allow sequential adhesion to different cell types during natural infection⁵.

The chromosome of *N. gonorrhoeae* strain MS11 contains two expression loci, *pilE1* and *pilE2*, involved in production of pilus protein and located within about 26 kilobases of each other on the chromosome³. When cells are in a pilated state both these genetic loci carry intact pilin coding sequences as well as their own promoters. Phase variation from P⁺ to P[−] states involves deletion of pilus-coding sequences from one or both of the expressing loci². Deletions in *pilE* apparently involve single or multiple recombination events between directly repeated sequences that lie within these loci. Interestingly, deletion in only one *pilE* gene can result in reduced or no pilus expression despite the fact that the remaining gene is apparently intact and should be functional. This suggests that there is a further mechanism (possibly a *trans*-acting activator) for regulating pilus expression that is itself regulated by the recombinational switching events.

It is not yet clear whether pilin sequences deleted from *pilE* are destroyed, as occurs in yeast mating-type conversion, or whether they are parked in silent, non-expressing loci somewhere else on the gonococcal genome, as occurs in antigenic variation of surface glycoproteins among certain trypanosomes. Certainly the ability of a P[−] derivative to switch back to producing homologous pili suggests that the information necessary to produce a specific pilin is conserved. However, a functional expressing gene can be regenerated by recombination between a deleted *pilE1* and an intact *pilE2* locus². This may be part of the reason for the duplication of *pilE* in the gonococcal genome, although maximization of pilus expression and the potential to elicit two distinct pilins simultaneously may also be involved.

The chromosome of the gonococcus contains, in addition to the two *pilE* loci, several silent sequences (*pilS* loci) that are partially homologous to the expressed pilin genes but which lack their own promoters^{2,3}. One such silent locus, *pilS1*, lies within about 15 kilobases of *pilE1* and contains sequences that are not colinear with the mature pilin coding sequence.

Sequences corresponding to the variable part of pilin are over-represented in *pilS* by about 15–20 times compared with those encoding the constant part³. This suggests that the silent regions contain information necessary for generating antigenic diversity. Furthermore, regeneration of intact pilin-expressing loci during a P[−] to P⁺ phase change can be explained by gene-conversion events that would result from recombination between a *pilS* locus and a partially deleted *pilE* site^{1,2}.

What is not yet clear is whether *pilS1* or some other more remote silent locus is the active donor of DNA sequences corres-

ponding to the variable part of pilin. The enzymatic mechanism that effects the required intragenic recombination events is also unknown. It will be interesting to discover whether specific enzymes exist for this purpose or whether the normal generalized recombination system of the gonococcus is sufficient.

It should be remembered that pili are only one component of a complex armoury of virulence determinants expressed by the gonococcus. One such determinant is principal outer membrane protein II (PIL or Op), which also undergoes both a phase change and antigenic variation. Interestingly, *opa*, the structural gene for this protein, is located within 1 kilobase of *pilE1*, although the two genes do not seem to be coordinately regulated nor are gross genome rearrangements involved in the modulation of expression⁶.

The genetic mechanisms used by *N. gonorrhoeae* for alteration of pilus type have parallels in other parasites that exhibit antigenic variation and in processes such as the generation of immunoglobulin diversity. Much of the evolutionary success of this pathogen, which has no host other than man, may be attributed to its complex devices for ensuring variation in its major surface antigens. □

1. Hagblom, P., Segal, E., Billyard, E. & So, M. *Nature* **315**, 156 (1985).
2. Segal, E., Billyard, E., So, M., Storbach, S. & Meyer, T. *Cell* **40**, 293 (1985).
3. Meyer, T., Billyard, E., Haas, R., Storzback, S. & So, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6110 (1984).
4. Rothbard, J.B., Fernandez, R., Wang, L., Teng, N.N.H. & Schoolnik, G.K. *Proc. natn. Acad. Sci. U.S.A.* **82**, 915 (1985).
5. Virji, M. & Heckels, J.E. *J. gen. Microbiol.* **130**, 1089 (1984).
6. Stern, A., Nickel, P., Meyer, T. & So, M. *Cell* **37**, 447 (1984).

J.R. Saunders is in the Department of Microbiology, University of Liverpool, PO Box 147, Liverpool L69 3BX, UK.

Polar studies

History of an Antarctic lake

from David A. Peel

DESPITE its immense ice-cap, Antarctica is in fact a polar desert where extremes of cold and aridity have created several unusual features. Since the earliest days of exploration, scientists have been fascinated by more than 20 ice-free areas or oases, which are now known to occur where local weather conditions determine that the overall rate of ice sublimation from the surface exceeds the annual rate of snowfall. These areas are completely free of ice providing that local geographical barriers prevent ice flow from a neighbouring area where it is accumulating. Within these dry regions, perennially ice-covered lakes are fairly common, but in many cases satisfactory models of their histories and physical characteristics have yet to be developed. On page 131 of this issue, Hermichen *et al.*¹ present new isotope data from a large freshwater lake in East Antarctica and demonstrate that the present-day lake is the concentrated residue of a body of water originally more than 50 times larger.

The perennially ice-covered lakes of the Antarctic show some unusual chemical and physical properties². Normally highly saline, although basically of freshwater origin, they occur in areas where the mean annual air temperature is around minus 20°C. They are usually covered by a relatively thin layer of ice overlying water which is at least 20°C, and sometimes up to 50°C, warmer than their surroundings. As the occurrence of life depends on the presence of liquid water, the lakes have been a focus of interest for biologists seeking to understand the adaptation of organisms to extremes of temperature³. Although the general morphology of the lakes has been understood for some time, they are individually a result of often quite complex histories.

One basic physical characteristic that has a direct bearing on the biological productivity of a lake, the thickness of the ice cover, has been modelled recently⁴. It has long been known that for lakes that are not frozen to the bottom, the ice thickness is relatively constant from place to place despite wide variations in the temperature and chemistry of the water. More than two decades ago it was suggested⁵ that solar heating was probably the main source of energy for the Antarctic lakes, with a skin of exceptionally transparent ice at the surface of the lake acting rather like greenhouse glass. Each summer, melt water flows into the lakes beneath the ice, balancing water lost by evaporation from the ice on the surface. Under steady-state conditions, it has been suggested that each winter a thickness of ice freezes onto the base of the floating ice, generating latent heat which balances the loss of sensible heat through the ice cover.

McKay *et al.*⁴ have recently developed a simplified model of the energy balance at the surface of Antarctic lakes based on analysis of annual average heat fluxes. The thickness of the ice cover is apparently sufficient to damp out most of the seasonal variations in the fluxes of heat from sources and sinks. Using actual measurements of solar flux and albedo values from one of the best-studied lakes (Lake Vanda, southern Victoria Land), and direct measurements of the typical light-transmission characteristics of the ice, these authors have been able to predict the equilibrium ice thickness as a function of the surface ice ablation rate. They found that this rate of ablation is the prime control of ice thickness; measured ablation rates of around 30 cm per year lead to predicted equilibrium ice thicknesses in the range of

3–4 m, in good agreement with what is observed. Lakes at higher altitudes lie closer to the snow line and suffer less ablation; moreover, the ice is often anchored to the bottom and hence the latent heat contribution from refreezing melt water ceases to be an important factor. In these lakes much greater thicknesses of ice can be maintained under a given air-temperature regime.

Analysis of the stable isotope composition of lake waters has proved to be a valuable tool in deciphering the evolution of present-day water systems⁶. Lake Untersee, the subject of the study by Hermichen *et al.*¹, was first probed by a Soviet team in 1969. The lake has an area of about 10 km² and lies in an region of high ablation (about 20 cm per year). Refreezing of melt water supplied by the adjoining glacier leads to an equilibrium ice cover which is 2.5 m thick. It is difficult to envisage how such a large volume of fresh water could have accumulated under present-day climatic conditions (the average temperature is roughly minus 20°C). Both the stable isotope data and the lack of any evidence for modern tritium indicate that the water has originated from the melting of old glacier ice. In this case, the concentration of salts in the relatively saline water now found can be accounted for if the present-day body of water is the residue resulting from the evaporation of a volume originally more than 50 times larger. The new isotope data provide some additional support for this idea. The present-day water is isotopically lighter than the snow that currently falls in the catchment area of the lake. Furthermore, the composition with respect to the Meteoric Water Line (a steady relationship between $\delta^{18}\text{O}$ and $\delta^2\text{H}$ generally observed in atmospheric precipitation, which reflects mainly equilibrium fractionation of the isotopes) is displaced in the direction of preferential depletion of ^{18}O . Both results can be explained satisfactorily if it is assumed that the water has been lost from the lake by sublimation of ice from the upper surface. The isotope composition of the sublimating water is then determined by isotope exchange during the freezing of water where ^{18}O is preferentially deposited. The indications are that this process has been proceeding more or less continuously since a period of warmer climate in the early Holocene. There are signs of biological primary production in Lake Untersee; future studies of organic deposits on the lake bed may help to trace the lake's evolution. □

1. Hermichen, W.D., Kowski, P. & Wand, U. *Nature* **315**, 131 (1985).
2. Wilson, A.T. *Antarctic Res. Ser.* **33**, 185 (1981).
3. Heywood, R.B. in *Antarctic Ecology* (ed. Laws, R.M.) Vol. 1, 279 (1984).
4. McKay, C.P., Clow, G.D., Wharton, R.A. Jr & Squyres, S.W. *Nature* **313**, 561 (1985).
5. Wilson, A.T. & Wellman, H.W. *Nature* **196**, 1171 (1962).
6. Nakai, N., Kiyosu, Y., Wada, H. & Takimoto, M. in *Memoirs of National Institute of Polar Research* No. 4 (ed. Torii, T.) 30 (1975).

David A. Peel is at the British Antarctic Survey, National Environmental Research Council, Madingley Road, Cambridge CB3 0ET, UK.

Why should plants be evergreen?

SIR — I read with interest Peter D. Moores's article "Why be an evergreen?" (*Nature* 313, 703; 1984). One factor referred to as a possible explanation for leaf longevity was leaf fall as a potential sterilizing agent. It was stated that many evergreens grow in impoverished soils. It has been argued in the past that plants growing in these conditions need to invest more heavily in chemical protection against phytophagous insects since growth on poor soils is slow, increasing vulnerability to attack. Many evergreens appear to be particularly rich in chemicals which defend them from phytophagous insects (terpenes associated with pine, for example). On the other hand, where growth is less costly, the alternative defence against phytophagous insects may be simply to shed leaves synchronously — that is, to adopt the deciduous habit to cleanse the tree annually of pests.

This hypothesis is open to an experimental test. If an evergreen is fumigated to remove phytophagous insects at the same time a deciduous species is coming into leaf, phytophagous insect population growth, measured as insect biomass per tree biomass, should be more rapid on the latter than the former. A favourable result would, however, be indicative rather than conclusive support of the hypothesis.

The hypothesis would be suspect if the maximum insect biomass achieved per tree biomass were higher on the evergreen. The reverse of this result would be consistent with the hypothesis whereas similar maxima would be the least informative outcome (although not inconsistent with the hypothesis).

BARRIE PEARSON

13 Knoll Cottage Residential Park,
Winfrith,
Dorset DT2 8LD, UK

Assessing the risk of dioxin exposure

SIR — The health risk of exposure to 2,3,7,8-tetrachlorodibenzodioxin (dioxin) is a subject of considerable international interest which has been addressed recently in *Nature*¹ and *Science*^{2,3}. As noted many times, understanding the nature of risk to humans is difficult because of the relative lack of information from epidemiological studies. Because of this, great emphasis has been placed on the results of two papers drawing on the mortality experience of workers exposed to dioxin at the Monsanto chemical plant in Nitro, West Virginia^{4,5}.

As with many occupational health studies, there are problems with these papers in the definition, and consistency of definition, of exposure. Zack and Suskind⁴ report nine cancer deaths in the cohort of 121 workers exposed to tetrachlorodibenzodioxin in a specific trichlorophenol pro-

cess accident, occurring in 1949. These nine deaths (Table 2 in the paper) give details of the individual's year of birth, year of hire by Monsanto, year of death and the cause of death. In the second paper, by Zack and Gaffey, published three years later⁵, there are lists of death from cancer amongst workers at Monsanto who are described as having been exposed to the herbicide 2,4,5-T. Workers exposed to 2,4,5-T were also exposed to varying concentrations of dioxin and were classified on the basis of having contracted the sentinel indicator disease chloracne. Zack and Gaffey report 25 deaths from cancer in the controls, workers not exposed to 2,4,5-T. This is referred to in Table II of the paper⁵.

In comparing the two papers, some men are listed as exposed to dioxin in one paper, but as not exposed in the other. In Table 2 in the Zack-Suskind paper, four cases are described as exposed to dioxin, whereas in the Zack-Gaffey paper, Table II reports the same four being not exposed (to 2,4,5-T and dioxin). Cases 1, 2, 5 and 7 in the Zack-Suskind Table 2 are those listed as being exposed to dioxin whereas in Table II of the Zack-Gaffey study, the same individuals are described (lines 6, 7, 9 and 22 of table) as being non-exposed. Judith Zack, an author on both papers, has reported that these individuals are one and the same⁶.

In addition, a number of exposed individuals may not have been included in the second Monsanto study. In Monsanto records, there are some 19 individuals who died of circulatory disease or cancer whilst in employment at the company, and who meet the criteria for inclusion in the exposed group. No reason is given for not including these cases.

Determining exposure in the context of the Nitro plant is clearly difficult, since conditions may have produced general exposure to much of the work force. Those men exposed to the chemical in the 1949 trichlorophenol process accident are said to have been seriously exposed, on the basis of developing more severe chloracne, than the men exposed to dioxin in the course of making 2,4,5-T over several decades at Nitro. It may be the case that some of the more severe chloracne cases were the workers exposed to dioxin in the trichlorophenol process accident. However, there was a similar spectrum of chloracne incidence in both groups according to a study in 1953⁷. Of the 117 cases of chloracne determined in 1953 in individuals exposed to dioxin as a result of the 1949 accident, 10 workers had left the company, 80 were said to have recovered from the skin disease, 24 had a mild form of chloracne, and 3 had moderately severe disease. In other workers exposed during 2,4,5-T manufacture, 97 additional cases of chloracne were reported as having occurred, 51 of these cases were regarded as being clear of the disease, 36 were said to have a mild form and 3 to have moderately severe skin disease. Seven of the 97 men

had left the company. Thus the spectrum of chloracne in the two groups seems similar. Another reason why it is inappropriate to separate these two exposed groups comes from information from a former manager of the Nitro plant⁸. In his deposition to court, Durland reported that some of the same workers were involved both in the clean-up after the trichlorophenol reactor accident and in making 2,4,5-T.

It may not be appropriate to classify degrees of dioxin exposure solely on the basis of chloracne as an indicator of relatively acute high-dose exposure to certain chlorinated hydrocarbons, including dioxin⁹. Other health effects — notably serious diseases of long latency, such as cancer — may be expected to result from lower and more chronic exposure¹⁰⁻¹³. Using one outcome variable to predict others is not advisable. Occupational studies usually use work histories, biological monitoring, or environmental measurements to determine, or approximate, exposure of subjects.

In view of the importance of the Monsanto cohort for determining human health risks of dioxin, it is our view that data from the company need to be reassessed. The total cohort of workers exposed to dioxin at Monsanto should be considered as a whole without making a distinction, as apparently done by Suskind, Zack and Gaffey, between workers exposed to dioxin in the trichlorophenol process accident or when making 2,4,5-T.

It may be that the overall number of deaths — from cancer, circulatory disease or other causes — is no greater than an appropriate control cohort. Data from a recent study by Moses *et al.*¹⁴ on this same population suggests that there is an excess of cardiovascular-related deaths¹⁵.

The epidemiological picture at Monsanto remains confused. We believe that more information will only be established by re-examining all of the data, and, given international concerns over dioxin, that this needs to be done urgently.

ALASTAIR HAY

Department of Chemical Pathology,
University of Leeds,
Leeds LS2 9NL, UK

ELLEN SILBERGELD

Environmental Defense Fund,
1525 Eighteenth St. NW,
Washington DC 20036, USA

1. *Nature* 303, 104; 304, 6 (1983); 308, 102 (1984).

2. Abelson, P.H. *Science* 220, 1337 (1983).

3. Sterling, T.D. *Science* 223, 1202 (1984).

4. Zack, J.A. & Suskind, R.R. *J. occup. Med.* 22, 11 (1980).

5. Zack, J.A. & Gaffey, W.R. in *Human and Environmental Risks of Chlorinated Dioxins and Related Compounds* (eds Tucker, R.E., Young, A.L. & Gray, A.P.) 575 (Plenum, 1983).

6. Deposition of Dr Judith Zack US District Court for southern West Virginia. Civil Action N. 81-2098 I March 1984.

7. Atkin, R. *et al.* Report on a Clinical and Environmental Survey Monsanto Chemical Co. Nitro, W. VA by the Kettering Laboratory (Uni) Cincinnati, Ohio 1953).

8. Deposition of John R. Durland US District Court for southern West Virginia. Civil Action No. 81-2098 28 February 1984.

9. May, G. Br. J. ind. Med. 39, 128 (1973).
10. Kimbrough, R.D. (ed.) *Halogenated Biphenyls Terphenyls, Naphthalenes Dibenzodioxins and Related Products* (Elsevier, Amsterdam, 1980).
11. Hay, A. *The Chemical Scythe; Lessons of 2,4,5-T and Dioxin* (Plenum, New York, 1982).
12. Hardell, L. & Sandstorm, A. Br. J. Cancer 39, 711 (1979).
13. Hardell, L. Scand. J. Work env. Hlth 7, 119 (1981).
14. Moses, M., R. Willis, K.D. Crow, J. Thornton, A. Fischman, H. Anderson & Selikoff, I.J. Am. J. Ind. Med. 5, 161 (1984).
15. Moody, L. W.E. Halperin, M.A. Fingerhut & Landrigan, P.J. Am. J. ind. Med. 5, 157, (1984).

Dosage compensation and X-chromosome inactivation

SIR — It is now common practice to use the terms dosage compensation and X-chromosome inactivation interchangeably. It is argued here that this equation is technically incorrect and, more importantly, that it obscures a nontrivial facet of mammalian physiology. The proposition is that in flies and mammals, dosage compensation allows males to survive their monosomy for the X chromosome. In mammals, X inactivation is then a device that prevents females from expressing a tetrasomic level of X-chromosome gene product. That is, X inactivation allows females to survive dosage compensation. If we call X inactivation dosage compensation, we are overlooking something in males. That something is dosage compensation.

The idea of dosage compensation and its definition originated with H.J. Muller in his studies of *Drosophila*¹. In the context of the issues of his day, Muller saw dosage compensation from a selectionist perspective concerned with optimal phenotypes. In this view, dosage compensation was the mechanism by which selection has achieved equality of expression of sex-linked genes in single-X males and two-X females. In the modern context, however, we are forced to add an element to the Mullerian view: dosage compensation not only equilibrates male to female phenotypes but also — and perhaps primarily — equilibrates a male's single X to his two sets of autosomes. The point we seem to overlook is that a *Drosophila* male (or the heterogametic sex in any organism with a chromosomal sex determination system) is aneuploid and aneuploidy is deleterious.

We know from an extensive study of aneuploidy in *Drosophila*² that any heterozygous autosomal deficiency greater than about 3% of the haploid genome is a lethal condition. On the other hand, a *Drosophila* male is deficient for a whole X chromosome (about 20% of the genome). The issue, then, transcends his similarity to his sister. He requires dosage compensation for his survival. In *Drosophila melanogaster*, dosage compensation is effected by the elevation of the rate of transcription of the lone X in males relative to either of the X chromosomes in females³⁻⁵. More to the point, mutations in *Drosophila* which fail to hypertranscribe the X chromosome in males are male-specific lethals⁶.

Aneuploidy in mammals, as in *Drosophila*, is deleterious. Mammalian

males, like *Drosophila* males, have a congenital aneuploid condition for a considerable fraction of their genomes. We infer that dosage compensation in mammals, as in *Drosophila*, balances the expression of the single X chromosome in males to the autosomal complement; that is a male's single X is hypertranscribed relative to his autosomes. Hypertranscription in *Drosophila* males presents no problem for *Drosophila* females, as they transcribe their two X chromosomes at a basal rate balanced to the autosomes. However, in a species in which X hypertranscription is a property of all X chromosomes in either sex, the female has become, in a formal sense, a hyperploid. Since hyperploidy is also a deleterious condition, she would require a second system to achieve X to autosome balance. One option is to inactivate one of the two X chromosomes.

If mammalian X chromosomes are hypertranscribed in the interest of male survival, X-chromosome inactivation must be dosage-compensation compensation in the interest of female survival. This inference may be of some heuristic value in that it implies the existence in mammals of a class of genetic elements which function to modify X-chromosome expression and to which we might want to give some thought.

DAVID A. SMITH

Department of Biology,
University of California, San Diego,
La Jolla, California 92093, USA

1. Muller, H.J. Proc. 6th Int. Congr. Genet. (Ithaca) 1, 213-255 (1983); Harvey Lec. Ser. 43, 165-229 (1950).
2. Lindsley, D.L. et al. Genetics 71, 57-184 (1972).
3. Ananiev, E.V., Faizullin, L.Z. & Gvozdev, V.A. Chromosoma 45, 193-201 (1974).
4. Lucchesi, J.C. Am. Zool. 17, 685-693 (1977).
5. Lucchesi, J.C., Rawls J.M. & Maroni, G. Nature 248, 564-567 (1974).
6. Belote, J.M. & Lucchesi, J.C. Nature 285, 573-575 (1980).

High time for psychoimmunology

SIR — The term "psychoimmunology" was coined by Solomon¹ to describe his early studies of the effects of behaviour on immune function and on susceptibility to certain experimentally-induced disease processes. More recently, "psychoneuroimmunology"² has been used to refer to studies of the neuroendocrine mechanisms mediating the effects of behaviour on immune function — and vice versa. But in a News and Views article, "Psychoimmunology before its time", John Maddox questions the explanatory potential and, by implication, the heuristic value and practicality of pursuing such research³.

Despite an erroneous characterization of what is "psychosomatic", the article correctly implies that much of the impetus for studying interactions between the central nervous system (CNS) and the immune system derives from observations that psychosocial factors influence the development and progression of disease. To phrase this as a "familiar theme in literature",

however, is to impart a scientific triviality to such phenomena. In fact, there is a voluminous scientific literature that documents the integration of mind and body. To justify this scepticism, Maddox points out that the mechanism for such phenomena is unknown, but this does not mean it is unknowable — or that the phenomena are less real. In an attempt to account for such phenomena, then, it is quite reasonable to hypothesize that changes in immune function may mediate the effects of psychosocial factors on the development and/or progression of some pathophysiological states. Such a hypothesis is tenable, however, only if it can be shown that the CNS plays some role in the modulation of immunity.

In referring to a *kind* of connection that is made *plausible* by several unconnected observations, the article begrudgingly acknowledges a functional relationship between the CNS and the immune system. This is not the common knowledge that it is implied to be. That it is accepted at all is, in large part, a result of psychoneuroimmunological research conducted over the past 10 years. The scepticism expressed in the article, however, is not directed at the experimental findings of psychoneuroimmunology. It is fabricated, instead, out of unreferenced claims attributed to "psychoimmunologists": we know of no serious investigators who "talk as if there is no state of mind which is not faithfully reflected by a state of the immune system". The hypothesis that a behavioural state may have immunological consequences, however, cannot be dismissed as easily. After all, there are neurophysiological and neuroendocrine consequences to behaviour and accepting a link between the CNS and the immune system, it would be reasonable to expect that influences are exerted in the opposite direction. Analyses of the interactions between behaviour and the immune system are not a traditional part of either immunology or the behavioural or neurosciences, but perhaps they should be.

Space does not permit us to do more than point to the article's superficial analysis of the studies by Laudenslager *et al.*⁴ and Schleifer *et al.*⁵, the evaluations of research based on the time consumed and its cost, the false analogy between the (unreferenced) claims of psychoimmunologists and attempts to link states of mind with the development of cancer, or the difficulty of conducting psychoneuroimmunological research and the implication that the results obtained are neither reproducible nor significant. The basic flaw in the argument is revealed by what is claimed to be a crucial unanswered question: if a person's affective response to some event is capable of influencing immune function, why, the article asks, would such a person be more likely to die of a heart attack or even an automobile accident?

This question is not crucial; it is not even relevant. The relationship between an affective state and immune function may be

quite independent of the relationship between that affective state and other physiological processes and, in any case, depends upon a myriad of other biological and environmental factors. Furthermore, such a question could follow only from the underlying assumption that behaviourally-induced alterations in immune function represent the single mechanism by which psychosocial factors influence disease. We know of no psychosomaticist or psycho-immunologist who would accept such a proposition and we cannot imagine how such a hypothesis could be derived from the available literature.

There are probably those in every field whose exuberance outdistances their data. One is not required to share their exuberance, but one is obliged to consider their data. One can only wonder, then, why anyone would want to throw out the baby with the bathwater. Connections between the CNS and the immune system are now being uncovered. Whatever forces were operating to set immunology apart, recent data suggest that much could be learned by studying immunoregulation as part of an integrated network of adaptive processes including behaviour. If not now, when?

ROBERT ADER

NICHOLAS COHEN

*Departments of Psychiatry and
Microbiology (Immunology),
University of Rochester School of
Medicine and Dentistry,
Rochester, New York 14642, USA*

1. Solomon, G.F. & Amkraut, A.A. *A. Rev. Microbiol.* **35**, 155-184 (1981).
2. Ader, R. *Psychosom. Med.* **42**, 307-321 (1980).
3. Maddox, J. *Nature* **309**, 400 (1984).
4. Laudenslager, M.L., Ryan, S.M., Drugan, R.C., Hyson, R.L. & Maier, S.F. *Science* **221**, 568-570 (1983).
5. Schleifer, S.J., Kellner, S.E., Camerino, M., Thornton, J.C. & Stein, M. *J. Am. med. Ass.* **250**, 374-377 (1983).

JOHN MADDOX WRITES — My article was a protest against those psychoimmunologists who attribute as yet unwarranted importance to the immune system (as distinct from other physiological systems) among the physiological correlates of behaviour. My statement that "people may be driven by adversity into decline is a familiar theme in literature" was not intended to demean psychoimmunology but, rather, to illustrate that the phenomena it seeks to account for have long been recognized.

My reference to heart attack and road accidents rather than, say, infection among the causes of death of those who have been psychologically shocked was intended to raise the precise question Ader and Cohen answered by saying that physiological systems other than the immune system may be involved.

I see nothing objectionable in the moderate statement of Ader and Cohen. Specifically, I can readily accept that there are many immunological consequences of, say, bereavement, although I do not believe that the tendency towards untimely death among bereaved spouses can be attributed primarily (or even at all) to the immune

system (which is not to say that that may not be an attractive speculation).

Underlying this dispute (which seems to me not to be a substantial disagreement) is an important question. These reductionist days, it is proper to seek to identify the mechanisms by which states of mind cause physical illness. The immune system is one candidate, perhaps the most obvious after the adrenal hormone complex. But we all know that Freud's place in intellectual life rests in part on his careful documentation of the phenomena of psychosomatic illness, and on his explanation in terms of the unconscious, nowadays a somewhat old-fashioned concept.

So what if further investigation should show that the nervous system, and especially that part of it within the skull, plays an important part in the causation of psychosomatic illness? Such a development would no doubt be hailed as an important step towards Freud's rehabilitation, as a prophet if nothing more.

The point of this speculation, in the present context, is not to say that that is how it will turn out, but simply to suggest that the less obvious fields for explanation may ultimately be more productive. Much of the abuse that reductionists have recently attracted stems from their weakness for the kinds of explanations which happen to be at hand. □

An unexpected result in classical electrostatics

SIR — A recent discussion of the structure of snowflakes (W.S. Mortley, *Nature* **313**, 638; 1985) mentions the possible role of electric forces. In this and similar problems, establishing the least energy configuration appears vital. I would like to draw attention to a surprising result, which, to my best knowledge, has not been mentioned earlier.

What will be a minimum-energy configuration of a system of N equal point charges q placed inside a circle?

At first glance the answer is obvious: the Coulomb repulsion will arrange all N charges at the vertices of a regular polygon inscribed into the circle (configuration A). This statement seems to be so obvious that, apparently, nobody bothered to verify it by a direct calculation of the total electrostatic energy of the system (W). However, such calculation, which can be easily performed on a programmable pocket calculator, shows, rather surprisingly, that this is true only for $N < 12$. In units q^2/R the total Coulomb energy W is equal to 48.57568 for $N = 11$ and 59.80736 for $N = 12$. An alternative configuration (B), which is also compatible with the total symmetry of the system, has $N - 1$ charges equally spaced along the circumference and one 'extra' charge at the centre of the circle. For $N = 11$, $W(B) = 48.62450$ (that is, greater than $W(A)$), but for $N = 12$, $W(B) = 59.57568$, that is, less than $W(A)$. The non-equality $W(B) < W(A)$ also holds for $N > 12$ (I verified this numerically to $N = 400$, but

very likely it is true for any $N > 11$).

The above result means that starting from $N = 12$, the second configuration B (with one charge expelled to the centre of the circle) takes over from A as the minimum-energy configuration. In other words, it will be energetically beneficial for the system of N equal point charges confined in a circle to 'expel' one charge to the centre of the circle (if $N > 11$). Similar effect (the 'spontaneous ejection' of one charge to the geometrical centre) will also be likely to happen for N charges inside the spherical surface ($N > ?$). This may lead to some modification of usual theorems of electrostatic stability which claim that at the state of the equilibrium all charges in the conducting body are always located on the surface. The above example clearly illustrates that this is not necessarily so.

ALEXANDER A. BEREZIN

*Department of Engineering Physics,
McMaster University, Hamilton,
Ontario, Canada L8S 4M1*

Spider pheromone 'packets'?

SIR — The scanning electron micrograph (SEM) of the pustuliform organ of the spider *Holcolaetis vidua* Lessert (Salticidae) reproduced in *Nature* of 7 March (p.17) shows spherical bodies on the surface. It is suggested that the spheres may be 'packets' of pheromone secreted by the pores, on the assumption that the organ is pheromone producing¹.

In reality the SEM shows that the spheres are not associated with the pores but occur on the pustuliform organs in proportion to the surface area occupied by the organs, about 25%. The implied low volatility of the spheres seems inconsistent with pheromonal function. The mean diameter of the spheres of 1.3 μm , size distribution, aggregation and general appearance are consistent with the spheres coating the chorion of spiders' eggs of all families, including salticids, that I have examined². These spheres may be found on the surface of spiders newly emerged from egg cocoons² and the brood guarding behaviour of salticids may expose the adults to such contamination.

W.F. HUMPHREYS

*Western Australian Museum,
Francis Street,
Perth, Western Australia 6000*

1. Wanless, F.R. 1984. *Bull. Br. Mus. (Nat. Hist.) Zoology series* **47**, 444-481.
2. Humphreys, W.F. *J. Zool. Lond.* **200**, 303-316 (1983).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a rather technical character which are not provoked by articles or letters previously published (where Matters Arising remains appropriate).

A multichannel seismic study of lithospheric flexure across the Hawaiian–Emperor seamount chain

A. B. Watts*, U. S. ten Brink*, P. Buhl & T. M. Brocher†

Lamont-Doherty Geological Observatory and * Department of Geological Sciences, Columbia University, Palisades, New York 10964, USA

† Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

Existing models assume that the thickened crust beneath seamounts is the result of a surface volcanic load flexing an elastic plate. New results suggest that flexed oceanic crust beneath the Hawaiian–Emperor seamount chain is underlain by a 4-km thick deep crustal body. We interpret the body as a deep crustal sill complex associated with the tholeiitic stage of volcano building along the chain.

PLATE tectonics is based on the concept of a strong rigid lithosphere that overlies a weaker asthenosphere¹. The principal evidence for the existence of a strong lithosphere on long time-scales ($>10^6$ yr) comes from studies of the way in which it responds to surface loads such as late glacial lakes, river deltas³ and oceanic islands and seamounts⁴. One of the best-studied loads^{5–11} is the Hawaiian–Emperor seamount chain in the Central Pacific Ocean. This volcanic load is suitable for study because: (1) it is located in the interior of a plate away from the complexities of plate boundaries; (2) it is one of the largest loads ($\sim 10^{18}$ kg) on the Earth's surface; (3) its tectonic setting is reasonably well known, enabling the relationship between load age and the thermo-mechanical properties of the oceanic lithosphere to be examined.

It has been known since the earliest gravity measurements¹² that the Hawaiian Islands are associated with large-amplitude free-air gravity anomalies that cannot be explained by local models of isostasy and require some form of regional compensation to account for them. Venning Meinesz⁴ and Gunn⁵ showed that gravity data over the islands are consistent with the flexure model of isostasy. According to their model, the Hawaiian Islands represent a load on the surface of the oceanic lithosphere which responds in a similar manner to an elastic plate overlying a weak fluid substratum. Walcott⁷, Watts and Cochran⁸, Watts¹⁰ and Kunze¹¹ have shown, using this model, that the best fit to the gravity data is for an effective flexural rigidity of the oceanic plate in the range 5×10^{29} – 5×10^{30} dyn cm, which corresponds to an effective elastic thickness, T_e , in the range 17.8–38.3 km.

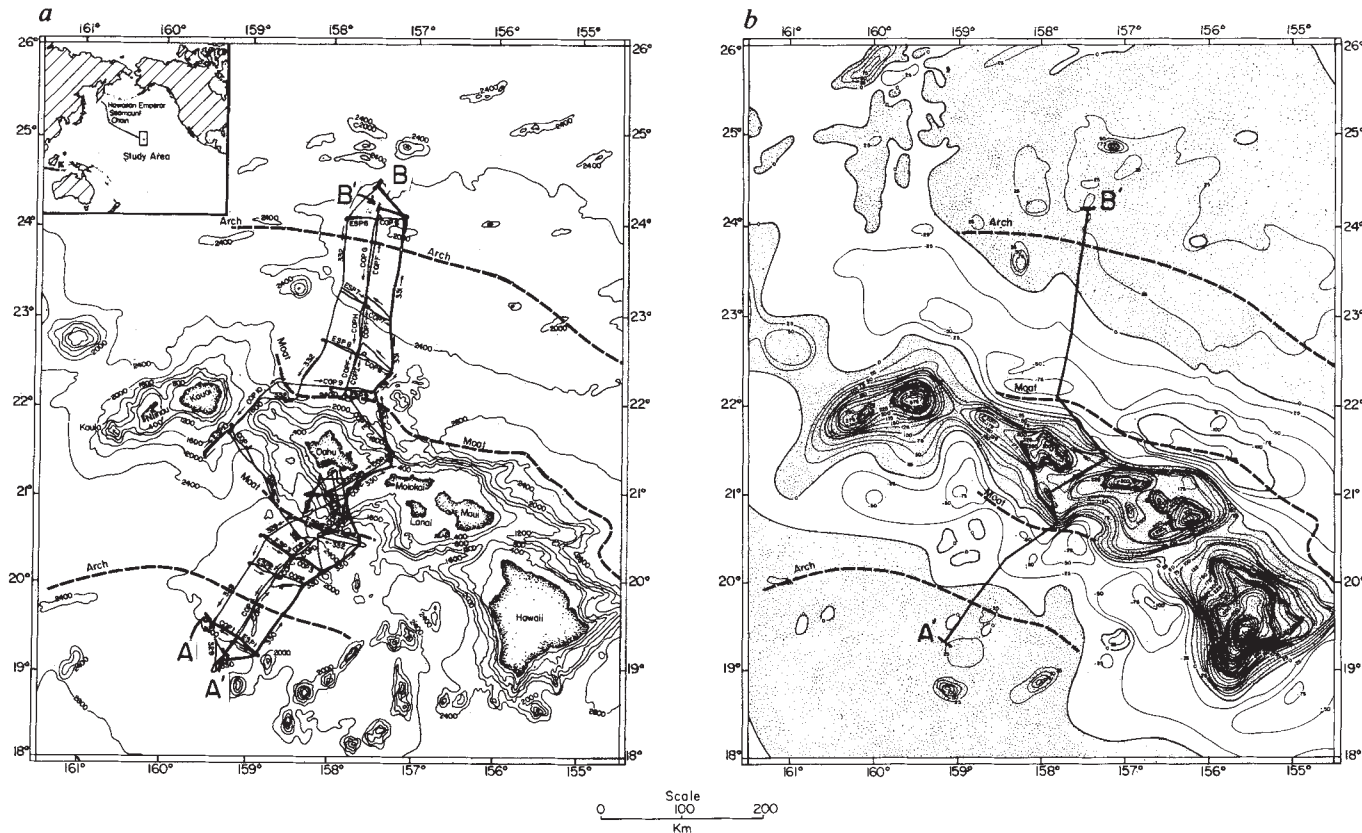


Fig. 1 Location map of the two-ship multichannel seismic experiment, *a*, Bathymetry; *b*, free-air gravity anomaly. In *a*, the contour interval is 400 fathoms (1 fathom = 1.8288 m) and is based on Chase *et al.*³⁴. Fine track lines, location of individual CDP, COP and ESPs; heavy line, location of ESP 1/COP 1 (Fig. 3) and CDP 330/CDP 331 (Figs 1, 2, profile AB); A'B', end points of the free-air gravity anomaly and bathymetry profile connecting ESP mid-points. In *b*, the contour interval is 25 mgal (1 mgal = 0.001 cm s⁻²) and is based on Watts and Talwani³⁵. Gravity anomalies >25 mgal are stippled. Thick dashed lines in *a* and *b* indicate the location of the Hawaiian arch and moat based on all available bathymetry profiles in the region.

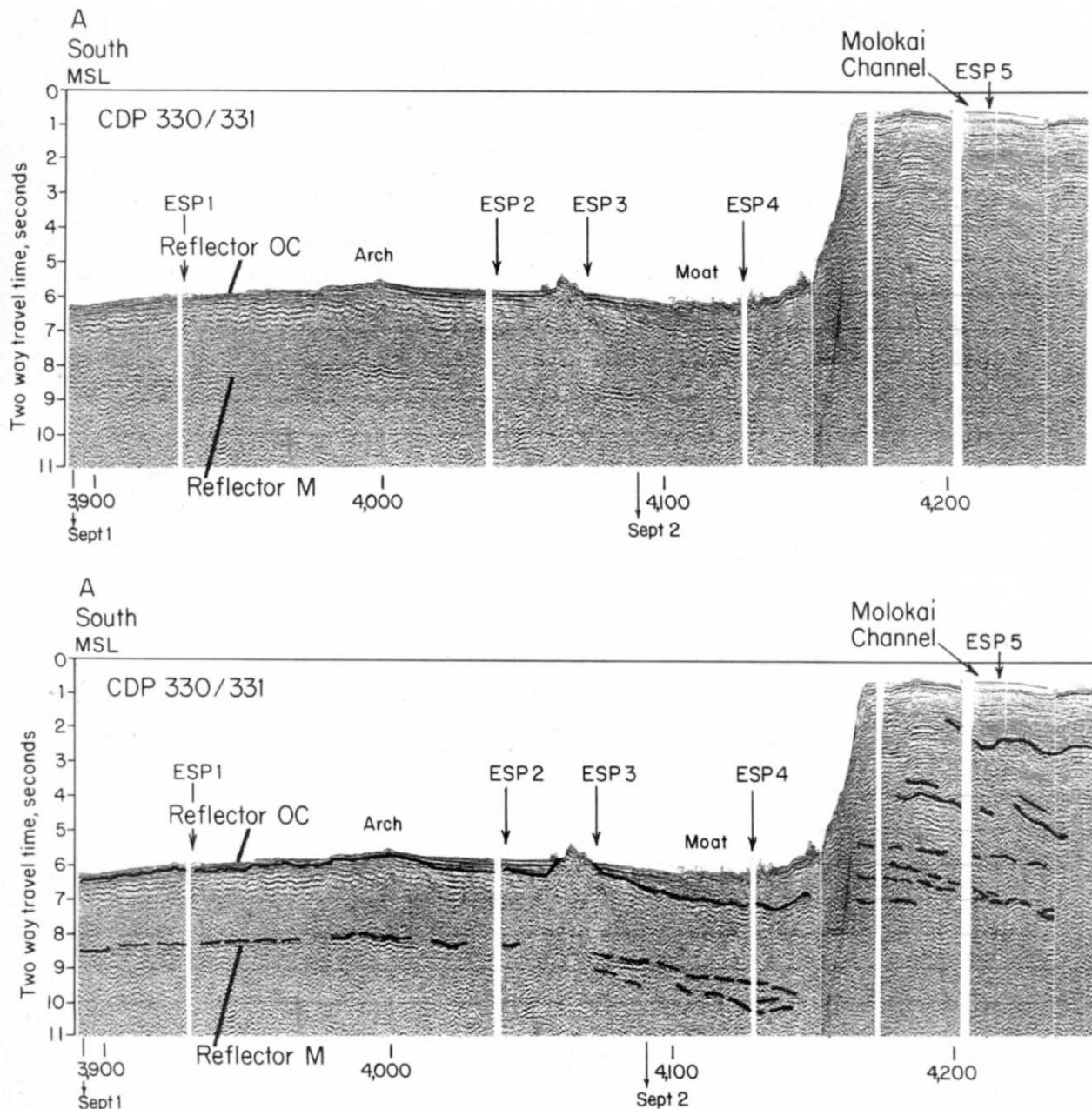


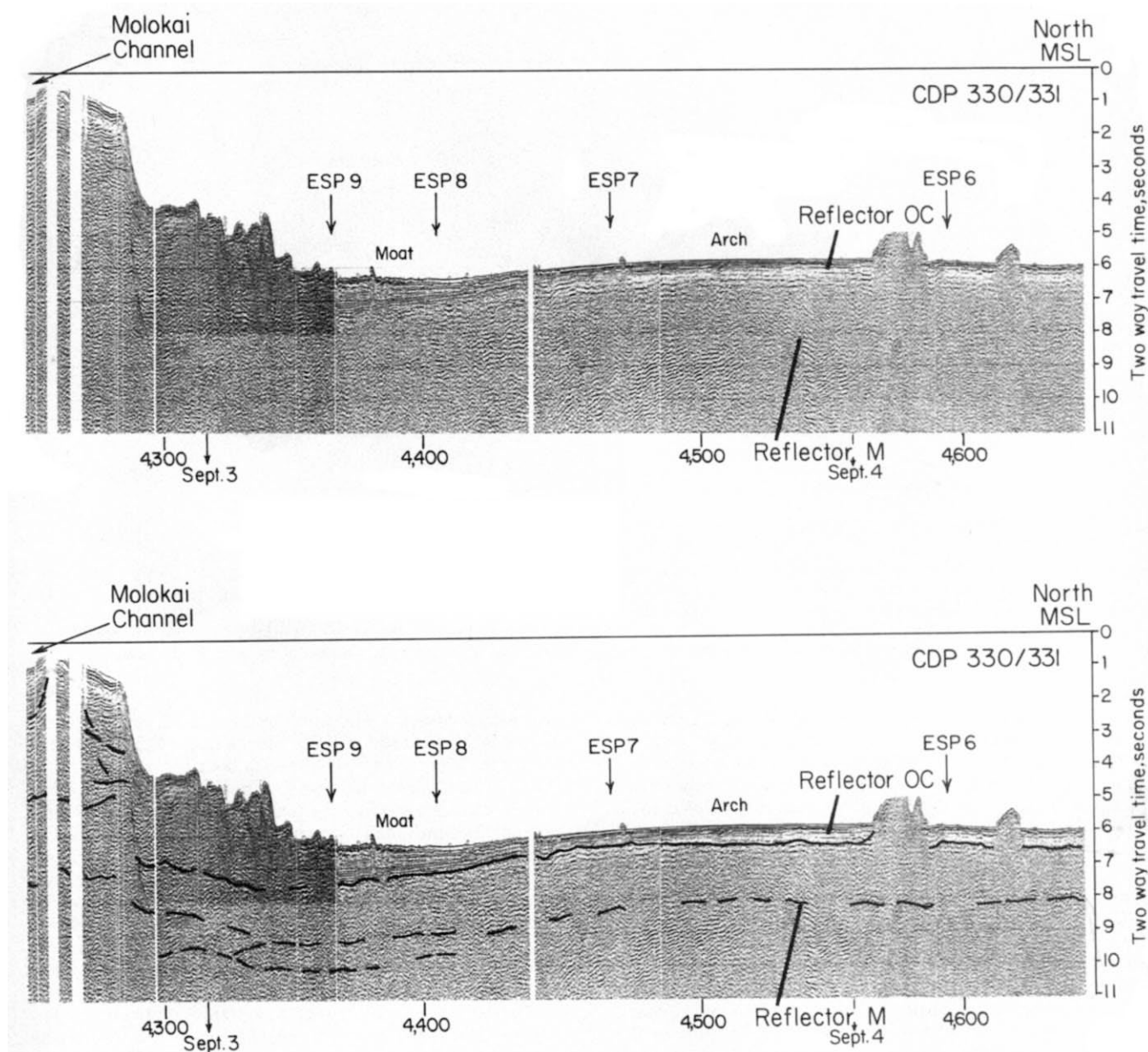
Fig. 2 (Above and opposite) CDP profiles 330 and 331 of the Hawaiian ridge and vicinity (Fig. 1a, profile AB). The 48-fold stacked section was acquired with a 3.6-km towed array. The black line on the upper section highlights prominent seismic reflectors. Reflector M at ~8.2 s below the arch is associated with the *M*-discontinuity and is continuous on a scale of ~50 km. Reflector OC is associated with the top of the oceanic crust. The number below the profiles is the accumulated distance in km along the ship track.

The actual value of T_c depends, however, on the type of flexural model that is assumed. For example, Walcott's⁷ estimate is based on a model that assumes the plate is fractured beneath the islands, whereas Watts and Cochran⁸, Watts¹⁰ and Kunze¹¹ derive their estimates assuming a continuous plate. Both types of model are in isostatic equilibrium and produce similar shape gravity anomalies to observed profiles. Thus, gravity data have not been able to discriminate unequivocally between different flexure models.

An independent test of flexure is to use seismic data to constrain the predicted deformation of the plate. For example, Walcott⁷ predicted that normal thickness oceanic crust is flexed by ~13 km beneath Oahu, whereas Watts¹⁰ and Kunze¹¹ predicted only ~4 km. If the thickness of the oceanic crust in undeformed regions is 6 km, then these authors imply a depth to the *M*-discontinuity of 23.5 and 14.5 km for the fractured and continuous plate models, respectively.

During the early 1960s, a number of seismic refraction experiments were performed near the Hawaiian Islands, principally in support of Project Mohole (ref. 13). These experiments showed that in the region of the Hawaiian arch the mean thickness of the oceanic crust is ~6.3 km, increasing to ~13.0 km beneath the moat and ridge. Subsequently, Furomotu *et al.*¹⁴ suggested that the crustal thickness is ~19–20 km beneath Oahu, thinning to ~17 km to the north-west beneath Kauai and ~10–14 km to the south-east beneath Hawaii. Unfortunately, these early seismic studies were hampered by poor navigation and drifting receivers and do not provide a suitable data set with which to constrain reliably the flexure model.

The purpose of this article is to summarize the results of a two-ship multichannel seismic experiment over the Hawaiian-Emperor seamount chain near the islands of Oahu and Molokai. A more detailed description of all the data obtained during the experiment will appear elsewhere (U. ten B. *et al.*, in prepar-



ation). The overall objective of the experiment was to define the detailed crustal and upper mantle velocity structure beneath a mid-plate volcanic island chain. We examine here the implications of the seismic experiment for the long-term ($>10^6$ yr) thermo-mechanical properties of the oceanic lithosphere and petrological models for the evolution of mid-plate oceanic volcanoes.

Seismic experiment

During August–September 1982, the Lamont–Doherty Geological Observatory and the Hawaii Institute of Geophysics carried out a two-ship seismic experiment near Oahu and Molokai in the Hawaiian Islands (Fig. 1). The experiment involved two vessels, the *Robert D. Conrad* and *Kana Keoki*, and used the single-ship multichannel seismic reflection profiling technique¹⁵ to determine the configuration of individual crustal layers and the *M*-discontinuity and the two-ship technique¹⁶ to determine the detailed velocity structure of the crust and upper mantle. The *Conrad* was equipped with a 3.6-km long seismic engineering multichannel streamer with 48 active channels and a large-volume (2,500 in³) high-pressure sound source. *Kana Keoki* was equipped with a single-channel streamer, a large-volume (2,000 in³) high-pressure sound source and 210 60-pound Tovex

charges. Navigation for each ship was by satellite. During the two-ship portion of the experiment, the range between the two vessels was continuously measured by Miniranger and Raydist systems.

The seismic experiment consisted of normal incidence common depth point (CDP) data, constant offset profiles (COPs) and eleven expanding spread profiles (ESPs) obtained across a 600-km long, 100 km wide 'transect' of the Hawaiian ridge centred near Oahu (Fig. 1). The first part of the experiment included a continuous 250-km long COP between the Molokai channel and the Hawaiian arch south of the Hawaiian Islands (Fig. 1a). While shooting the COP, the two vessels maintained a constant distance apart of 3.6 km and *Conrad* and *Kana Keoki* alternately fired their airguns each minute. At the end of the COP, the vessels re-grouped for an ESP (Fig. 1a, ESP 1). During the ESP, the two vessels separated from each other at a speed of ~ 5 knots while *Kana Keoki* fired its large airguns each minute and *Conrad* received. At the end of the 65–90-km-long line, the two vessels turned and moved toward each other along the same profile. The ESP was then reshoot with the *Kana Keoki* firing Tovex charges on a 10-min schedule. During the ESP re-shoot, *Conrad* fired its airguns each minute (with each 10-min shot missing) so that a CDP profile was obtained along the ESP. On

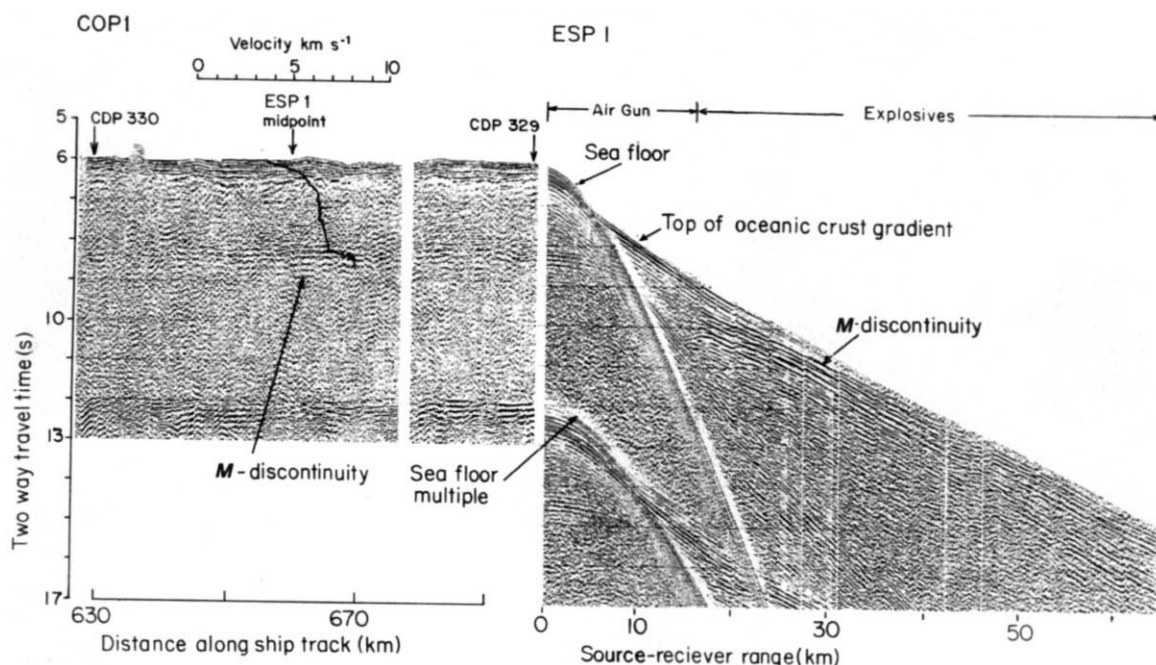


Fig. 3 ESP 1 and COP 1 of the Hawaiian arch, south of the Hawaiian ridge (Fig. 1a). The ESP data were obtained as the two ships separated at a constant speed from a common mid-point. At the end-point, the two ships steamed towards each other. During the 'outgoing' part of the experiment *Kana Keoki* fired its air guns and the *Conrad* received. During the 'ingoing' part *Kana Keoki* fired explosives. The ESP data shows a combination of the air gun and explosive record. The COP data were obtained by one ship (*Conrad*) firing its air guns and receiving. The black line on COP 1 represents the best-fitting velocity depth profile, obtained by a τ -p sum of the ESP X - T data.

completion of ESP 1, the vessels regrouped for a COP at a constant separation of 16 km to ESP 2. The COP/ESP sequence was then repeated until the transect had been completed south of the islands. The second part of the experiment was comprised of a similar sequence of COP/ESPs between the Molokai channel and the arch to the north. On completion of ESP 11, *Conrad* began a 1,200-km long CDP profile connecting the ESP end-points. During this portion of the experiment, *Conrad* fired its airguns on a 20-s schedule.

The data acquired during the seismic experiment were processed using techniques described by Stoffa *et al.*¹⁷. The ESPs were corrected for ship's navigation, the ranges between shot and receiver, shot times and clock synchronization. After demuxing the data, individual seismic traces were sorted and summed in bins of 50 and 100 m. We performed τ - p mapping of the stacked X - T data and carried out a τ -sum inversion (assuming horizontal homogeneous layers) on the τ - p data to yield a velocity-depth function at each ESP mid-point¹⁷. The solutions from τ -sum inversion were verified by forward modelling of first arrivals using velocity gradients and accounting for spherical losses. Also, the velocity-depth function was compared with the results of semblance analysis of CDP data near some ESP mid-points and the two-way travel time data on the CDP profile that crossed each ESP mid-point. The CDP data were processed using routine methods with some exceptions. Before stacking, velocity filtering was performed on parts of the profiles, each CDP gather was muted and the first and last traces eliminated. This additional processing removed multiples of the sea floor and sub-bottom reflectors and reduced the scattered energy associated with irregularities in each reflector. After stacking and predictive deconvolution, time-varying band pass filtering was applied to all the CDP data.

Results

We present here selected profiles of the ESP and CDP data gathered during the seismic experiment. Figure 2 shows CDP profiles 330 and 331 obtained by *Conrad* on a transect of the Hawaiian ridge between Oahu and Molokai (Fig. 1a). The figure shows a strong reflection (reflector M) from the M -discontinuity at ~ 8.2 s two-way travel time beneath the Hawaiian arch south

of the islands and at ~ 8.0 s north of the islands. In general, the reflector is stronger south of the islands than to the north. The travel time difference between the sea floor and the reflector is ~ 2.0 s north of the islands and 2.4 s to the south. We interpret these differences as indicating that the oceanic crust is thicker south of the islands than to the north. The travel-time difference between the sea floor and reflector M increases from the arch to the moat indicating an increase in the thickness of the crust beneath the moat. At $\sim 4,070$ km (Fig. 2), there is evidence that reflector M 'splits' at a two-way travel time of ~ 9.1 s. We interpret this split in the M -discontinuity as the result of a deep crustal layer underlying the oceanic crust beneath the moat and ridge. The top of the oceanic crust is characterized by a generally weak reflector (OC) with rough topography underlain by low frequency 'multiples'. Other strong reflectors are associated with the poor-to-well stratified material of the moat infill.

Figure 3 shows ESP 1 (Fig. 1a) obtained by *Conrad* and *Kana Keoki* near the Hawaiian arch south of the islands. The right-hand side of the figure shows a line of refracted arrivals which are nearly tangential to the top of oceanic crust and M -discontinuity reflectors. There are departures, however, from straight refraction lines. For example, at ranges of 5–25 km there is a convex downward curvature of the primary refracted arrivals, suggesting a continuous increase in velocity with depth, while at ~ 27.5 km there is a gap in primary arrivals, suggesting an abrupt decrease in the velocity gradient with depth. The CDP profile (Fig. 3) shows a strong reflector associated with the M -discontinuity between 8.15 and 8.45 s. The solid line at the ESP 1 mid-point on the CDP is the velocity-depth function obtained by τ - p mapping of the X - T data and τ -sum inversion. The velocity-depth function shows a progressive increase from 2.5 to 8.1 km s^{-1} . Abrupt changes in the velocity gradient are associated with the top of the oceanic crust and the M -discontinuity.

The principal results of the CDP data and ESPs 1–9 are summarized in Fig. 4. The velocity-depth data at each ESP mid-point and the depth to prominent reflectors on CDP 330/331 have been projected onto bathymetry profile A'B' (Fig. 1a). Figure 4 shows that the total crustal thickness increases from ~ 6.5 to 6.8 km beneath the Hawaiian arch to ~ 10.5 –12.0 km

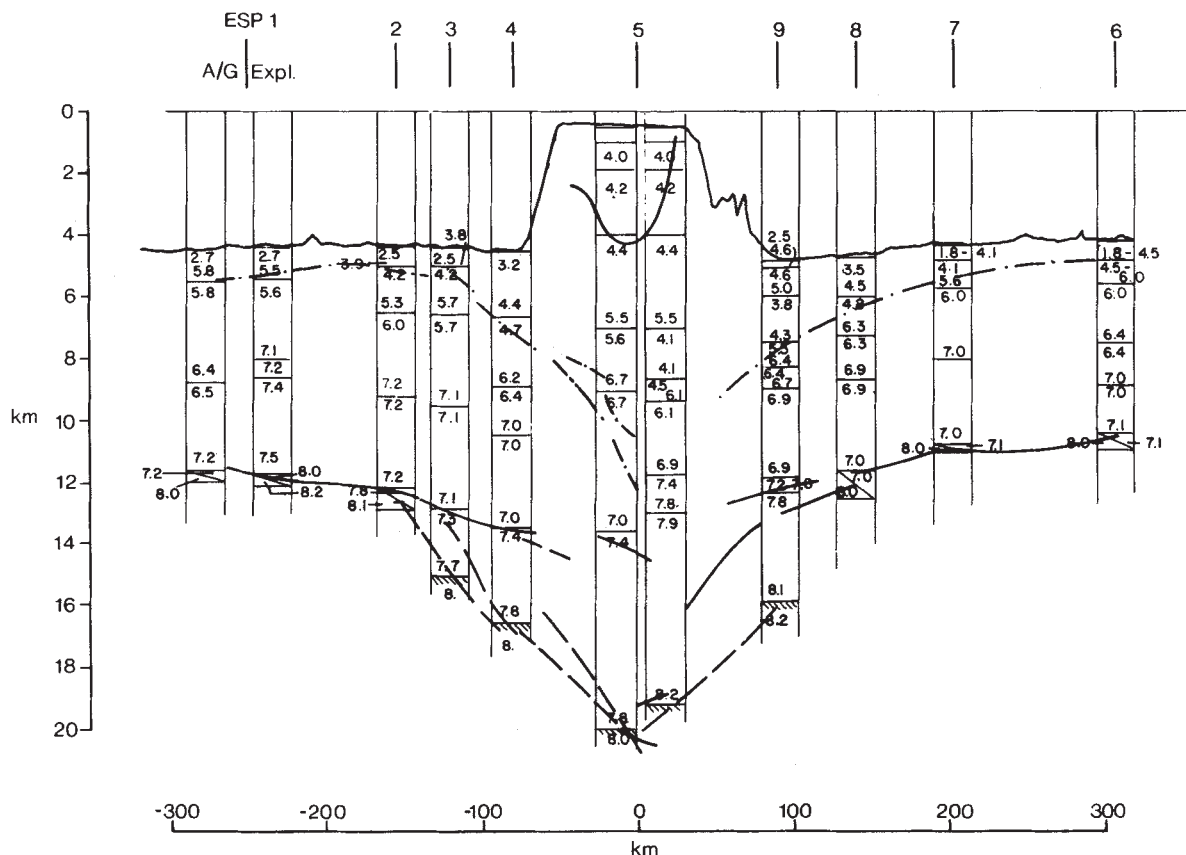


Fig. 4 Summary of the velocity and crustal structure obtained on ESPs 1-9. Horizontal lines in each column represents the depth to a prominent change in velocity with depth. The numbers refer to the velocity (km s^{-1}) immediately above and below the velocity change. The dashed-dot, continuous, and dashed lines represent prominent reflectors in Fig. 2 which have been converted to depth using the ESP velocity data results. These three reflectors are interpreted as corresponding to the top of the oceanic crust, the base of the pre-flexed oceanic crust and the *M*-discontinuity respectively. The ESP data have been plotted at the intersection of each ESP with profile A/B'. Note that two possible solutions are shown for ESP1 and ESP5.

beneath the moat, and 18.0–19.0 km beneath the ridge. In the region of the arch, the crustal layer below reflector OC has a velocity which grades from 4.2 to 5.6 km s^{-1} (ESP 1, 2, 6, 7) to ~8.0–8.2 km s^{-1} . Below reflector OC in the moat there is a sudden increase in velocity from 7.0 to 7.4 km s^{-1} at depths of 6.5 km below the sea floor (ESP 4) and from 6.9 to 7.2 km s^{-1} at 7.0 km to the north (ESP 9). At ESP 5 on the ridge, reflector OC cannot be as easily identified. Furthermore, the *X-T* data cannot be interpreted in terms of a single velocity-depth profile. Figure 4 shows two possible solutions for ESP 5, both showing a similar upper crustal structure with velocities increasing from 4.0 to 5.5 km s^{-1} . At greater depths, however, the two solutions differ. The solution, which does not include a low-velocity zone (LVZ) in the upper crust, is characterized by a sudden increase in velocity from 7.0 to 7.4 km s^{-1} at depths of ~13.0 km below the sea floor, whereas the model with a LVZ shows a sudden increase from 6.9 to 7.4 km s^{-1} at depths of ~11.3 km. Above reflector OC, velocities are generally in the range 2.3–4.0 s^{-1} . At ESP 9 in the moat region, there is evidence from arrivals on *X-T*, τ -*p* and from amplitude data for a LVZ above reflector OC. We interpret this LVZ as the result of a high-velocity lava flow capping low-velocity material of the moat infill.

To constrain better the crustal model, we computed the gravity and geoid effect of the seismic structure and compared it with the observed free-air gravity anomaly profile obtained on *Conrad* and *Kana Keoki* and the geoid anomaly derived from SEASAT radar altimeter data. Because of uncertainties in the seismic interpretation at ESP 5 (Fig. 4), two models for the deep crustal structure of the ridge were considered: one in which the oceanic crust is thinned beneath the ridge (Fig. 5, model A) and one in which the crust is of normal thickness (Fig. 5, model B). Model A is consistent with the solution without a LVZ in the upper crust, whereas model B agrees better with the solution with a LVZ. For the purposes of the gravity and geoid calculation, the

velocity-depth data near the arch and moat were generalized to an upper 4.2–6.0- km s^{-1} layer, a middle 6.0–7.1- km s^{-1} layer and a lower 7.4–7.8- km s^{-1} layer. Densities of 2.6, 3.0 and 3.15 g cm^{-3} , respectively, were assigned to these layers using the velocity-density relationships of Christensen and Salisbury¹⁸. We assigned a density of 2.3 g cm^{-3} for the material filling the moat, a density similar to the average density of vesicular Hawaiian flows¹⁹ and in the range of values normally chosen for sediments. At ESP 5, the volcanic 'core' of the ridge is associated with velocities of 4.0–6.7 km s^{-1} . A density of 2.5–2.7 g cm^{-3} was assigned for the core of the ridge, similar to values normally considered for volcanic flows¹⁹.

Figure 5 shows that there is a good overall agreement between the observed and the calculated gravity and geoid anomalies based on the seismic structure. Gravity and geoid data agree with the suggested thickening of the crust from ~6.5 km beneath the arch to 12.0 km beneath the moat and 18.0–19.0 km beneath the ridge. Furthermore, the data are consistent with the existence of a deep crustal layer underlying oceanic crust beneath the ridge and a portion of the flanking moat. The data cannot, however, distinguish whether the oceanic crust is thinned beneath the ridge or whether it has a similar thickness as beneath the moat and arch. Figure 5 shows that if oceanic crust is normal thickness beneath the ridge a high density body is required beneath it to compensate for the additional low density associated with the increase in thickness of the oceanic crust. The origin of the body is not clear, but gravity modelling suggests that it represents a deep seated extension of the dense volcanic pipe associated with the Koolau caldera on Oahu²⁰.

Implications

The seismically determined crustal structure of the Hawaiian ridge has implications for both the thermo-mechanical

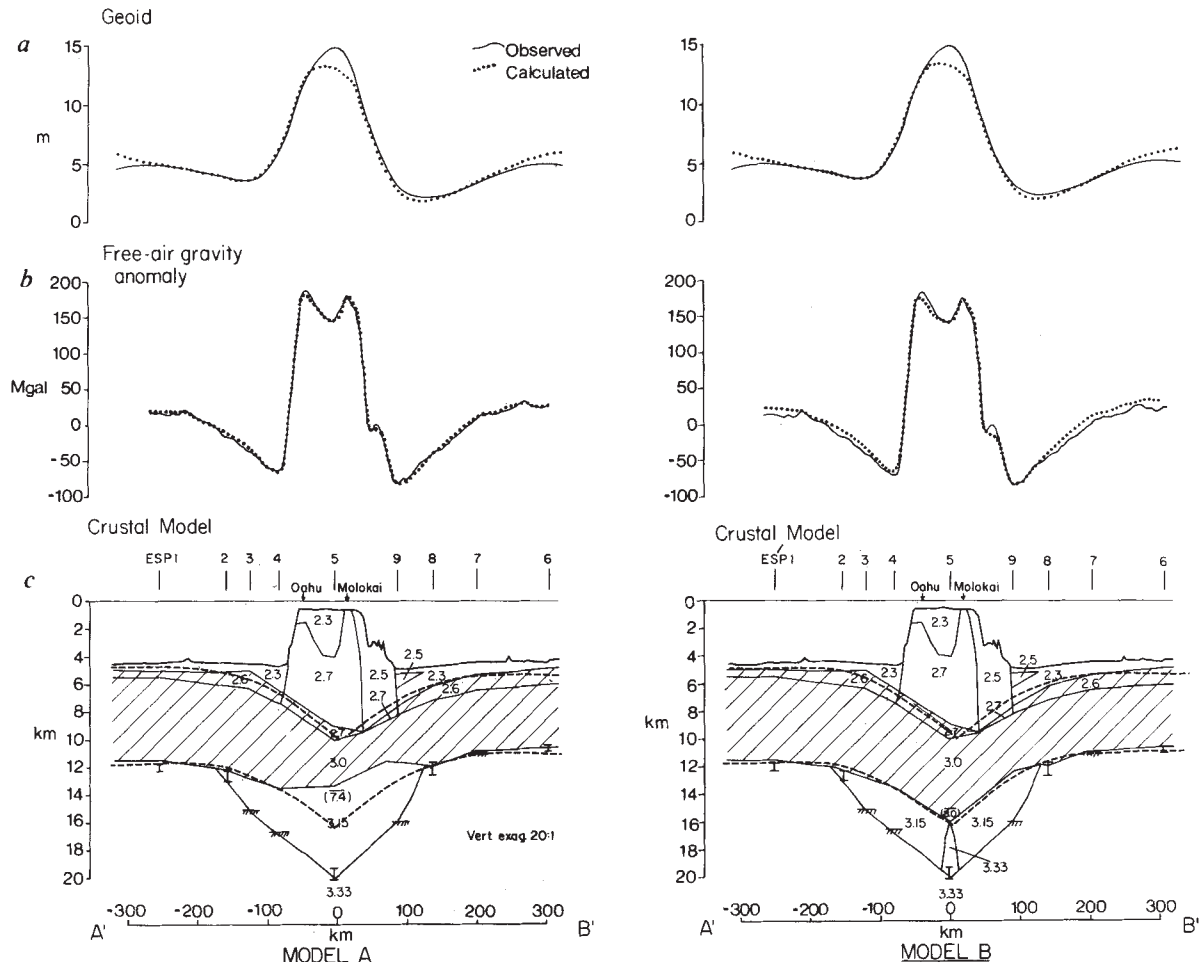


Fig. 5 Comparison of calculated geoid and gravity anomaly based on a density model constrained by seismic data to observed satellite derived geoid and surface-ship free-air gravity anomaly data. *a*, Geoid anomaly obtained by subtracting the geoid derived from a descending track of SEASAT from the GEM 10 Earth model complete to degree and order 12. *b*, Free-air gravity anomaly profile derived from data obtained by Conrad and Kana Keoki during COP A/COP B and COP E/COP F. *c*, Density model based on velocity-depth data in ESPs 1–9 and the velocity–density relationships of Christensen and Salisbury¹⁹. The configuration of the models (see text) have been adjusted so as to best-fit the gravity and geoid data and still explain the seismic data. Model A, thinned oceanic crust; model B, normal thickness oceanic crust. Heavy dashed lines are calculated curves based on a flexure model of isostasy with $T_e = 25$ km.

properties of oceanic lithosphere and petrological models for the evolution of mid-plate seamount chains.

The configuration of the oceanic crust near Oahu and Molokai is compared with predictions based on a flexure model of isostasy described in Fig. 5. In the model, it is assumed that the Hawaiian ridge and its associated infill material represent a two-dimensional load on the surface of an elastic plate overlying a weak fluid substratum. Thus, we assume that the only loads that act on the plate are surface loads and that buried loads²¹ do not contribute to the flexure. Figure 5 shows that there is a reasonable agreement between the configuration of the upper surface of the oceanic crust and the calculated flexure profile for $T_e = 25$ km. There is also good agreement between the calculated profile and the configuration of the base of the oceanic crust for the model in which the oceanic crust is uniform in thickness beneath the ridge. Figure 5 shows, however, that the agreement is poor for the model in which the oceanic crust thins beneath the ridge. These results, therefore, suggest that beneath the ridge, flexed normal thickness oceanic crust is either underlain by a high density body (model B) or has been altered by densification in its lowest part (model A).

The flexure profile in Fig. 5 is based on a single value of T_e and a fractured plate model. Other values of T_e would be required if different types of flexure model are used. For example, if a continuous rather than a fractured plate had been used, then a lower value of T_e (~ 20 km) would be required. We prefer the fractured plate model (and hence the higher T_e value) because it best explains the shape of the top of the oceanic

crust below the moat and the edge of the load. We caution against a too literal interpretation of the T_e value in Fig. 5, however, because it may be possible to construct other models (for example, a model in which T_e decreases gradually from beneath the arch to the ridge) that could explain the seismic, gravity and geoid data (Fig. 5) equally well. A value of $T_e = 25$ km can, therefore, be regarded as a 'best fit' to the regional crustal structure of the ridge and hence, a measure of the average response of oceanic lithosphere to volcanic island loading along the Hawaiian–Emperor seamount chain.

Oceanic flexure studies¹⁰ have shown that T_e is a strong function of the age of the lithosphere at the time of loading. Thus, volcanic loads formed on young lithosphere are associated with small values of T_e , whereas loads formed on old lithosphere are associated with large values. There is a good agreement between T_e and the depth to the 450 °C oceanic isotherm based on the cooling plate model. Bodine *et al.*²² have considered these results in terms of the rheology of oceanic lithosphere based on data from experimental rock mechanics and shown that

$$T_e \approx 4.3t^{1/2}, \text{ 'dry' olivine rheology} \quad (1)$$

$$T_e \approx 3.3t^{1/2}, \text{ 'wet' olivine rheology} \quad (2)$$

where t is age of the oceanic lithosphere (Myr) at the time of load emplacement. Using $t = 80$ Myr in equations (1) and (2) gives a predicted T_e of 30–40 km, which is somewhat higher than the value of T_e deduced here.

Although T_e at a volcanic load is a function for the thermal age of the underlying oceanic lithosphere¹⁰, it may not represent the actual temperature structure of the lithosphere because of perturbing effects associated with magmatic processes and/or subsequent thermal events. Crough²³, for example, suggested that the Hawaiian swell is the result of thermal uplift following reheating of oceanic lithosphere by the Hawaiian hotspot. He estimated from bathymetric data that the 80 Myr lithosphere near Hawaii has been reheated to a thermal age of ~30 Myr. Substituting this value in equations (1) and (2) gives a predicted T_e of 18–23 km, which is similar to T_e estimated here. However, Sandwell²⁴ computed the response of a moving plate over heat sources with different shapes and has shown that the peak in the temperature distribution may be displaced ≤ 20 Myr downstream of the source. The age of the volcanic load near Oahu and Molokai is only ~2–3 Myr, so that the T_e estimated here cannot be used to constrain the reheating hypothesis. Thus, the strongest evidence for the swell remains the exponential form of its subsidence.

Although the extent of lithospheric reheating is uncertain, it is clear that a flexure model of isostasy with $T_e = 25$ km can explain the overall configuration of the oceanic crust near Oahu and Molokai. The model cannot, however, account for the 7.4–7.8 km s⁻¹ velocity layer which seismic, gravity and geoid data suggest underlies flexed oceanic crust beneath the islands. This deep crustal layer is ~4 km thick, 200 km wide and generally symmetric below the centre of the ridge.

It is known from petrological studies that the Hawaiian volcanoes erupt lavas of different and distinct chemical compositions during their evolution. According to McDonald²⁵, the main stage of volcano building consists of large amounts of tholeiitic lavas which erupt rapidly (≤ 1 Myr; ref. 26). Following the main stage, the volcano usually erupts small amounts of alkali basalts. Finally, after a few Myr of quiescence, minor amounts of nephelinitic lavas may erupt from satellite vents. The exact nature of the source region of the tholeiites and alkali basalts is not known^{27,28} but seismicity studies near Hawaii²⁹ suggest that magma transfer associated with the tholeiitic stage of volcanism extends to depths of at least 50–60 km. The nephelinitic lavas are considered to be more primitive and to have originated from an even deeper source.

Seismic studies provide constraints on the velocity and density structure of the crust and upper mantle beneath a volcano and hence are of importance in developing petrological models. For example, seismic data³⁰ suggest that Mount Etna is underlain by a deep seated magma chamber.

From a consideration of all the seismic data obtained during the experiment, we believe that the most probable interpretation of the deep crustal body underlying flexed oceanic crust beneath Oahu and Molokai is that it represents a deep crustal sill complex. In particular, the body represents 'frozen' intrusions of tholeiitic composition into upper mantle rocks that were trapped in their ascent from a deeper source during the late phase of

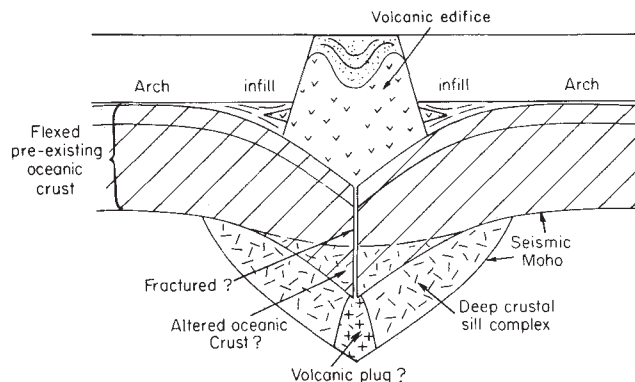


Fig. 6 Schematic model of the crustal structure of the Hawaiian ridge and vicinity based on the seismic, gravity and geoid data discussed here. The model has been generalized from models A and B in Fig. 5.

the tholeiitic stage or during the following period of quiescence. Such a model of a crust-mantle 'mix' agrees with seismic data showing that the deep crustal layer is intermediate in velocity between that of lower crust (gabbro) and upper mantle (peridotite) rocks. It is also in agreement with evidence^{31,32} which shows an absence of a deep crustal body below Kilauea and Mauna Loa on Hawaii. These volcanoes seem to be in the early phase of the tholeiitic stage and, therefore, would not be expected to be underlain by such a sill complex.

Figure 6 shows a schematic model for the deep crustal layer beneath Oahu and Molokai that explains the geophysical data and is consistent with current petrological knowledge. In proposing this model, we have eliminated (on density arguments) the possibilities that the deep layer represents either a magma chamber of basic picritic magmas which become ponded at or near the crust/mantle following their ascent through the mantle³³, or dunites and plagioclase-bearing pyroxenites that formed as early cumulates in a fractionating tholeiitic magma chamber²⁸.

The deep structure of oceanic islands and seamounts, however, is still too poorly known and other explanations may be proposed as more data become available. The seismic experiment described here was the first of its kind to be carried out over a mid-plate seamount chain. We believe that multichannel seismic studies of other seamounts offer particular promise of better understanding the origin of mid-plate volcanism in the future.

We thank J. Mutter, J. Glasser, C. Mountain, E. Vera, D. Lindwall, C. Zehnder and the officers and crew of RVs *Conrad* and *Kana Keoki* for their help at sea and J. Mutter and J. Weissel for critical reviews. This work has been supported by NSF grants OCE 81-11704 (L-DGO) and OCE 81-17210 (WHOI). Lamont-Doherty contribution 3797.

Received 7 December 1984; accepted 21 February 1985.

1. Barrell, J. J. *J. Geol.* **22**, 655–683 (1914).
2. Walcott, R. I. *Can. J. Earth Sci.* **7**, 716–727 (1970).
3. Cochran, J. R. *Bull. geol. Soc. Am.* **84**, 3249 (1973).
4. Venning Meinesz, F. A. *Gravity Expeditions at Sea 1923–1938*, Vol. 4 (Netherlands Geod. Comm. Delftse Uitgevers Maatschappij, Delft, 1948).
5. Gunn, R. *J. Franklin Inst.* **236**, 373–396 (1943).
6. Hamilton, E. L. *Bull. geol. Soc. Am.* **68**, 1011–1026 (1957).
7. Walcott, R. *Tectonophysics* **9**, 435–446 (1970).
8. Watts, A. B. & Cochran, J. R. *Geophys. J. R. astr. Soc.* **38**, 119–141 (1974).
9. Walcott, R. I. in *Geophysics of the Pacific Ocean Basin* AGU Monogr. **19**, 431–438 (1976).
10. Watts, A. B. *J. geophys. Res.* **83**, 5989–6004 (1978).
11. Kunze, A. W. G. *Tectonophysics* **69**, T1–T8 (1980).
12. Hayford, J. F. S. & Bowie, W. *Dept Commerce and Labor Spec. Publ.* **10** (Coast and Geodetic Service, Washington, D.C., 1912).
13. Shor, G. G. & Pollard, D. D. *J. geophys. Res.* **69**, 1627–1638 (1964).
14. Furumoto, A. S., Woollard, G. P., Campbell, J. F. & Hussong, D. M. in *The Crust and Upper Mantle in the Pacific Area* AGU Monogr. **12**, 94–111 (1968).
15. Mayne, W. H. *Geophysics* **27**, 927–938 (1962).
16. Stoffa, P. L. & Buhl, P. *J. geophys. Res.* **84**, 7645–7663 (1979).

17. Stoffa, P. L., Diebold, J. B. & Buhl, P. *Geophys. Prospect.* **30**, 25–57 (1982).
18. Christensen, N. I. & Salisbury, M. H. *Rev. Geophys. Space Phys.* **13**, 57–86 (1975).
19. Manghnani, M. H. & Woollard, G. P. in *The Crust and Upper Mantle in the Pacific Area*, 401–516 (American Geophysical Union, Washington, 1970).
20. Strange, W. E., Woollard, G. P. & Rose, J. C. *Pacif. Sci.* **19**, 381–389 (1965).
21. Karner, G. D. & Watts, A. B. *J. geophys. Res.* **88**, 10,449–10,477 (1983).
22. Bodine, J. H., Steckler, M. S. & Watts, A. B. *J. geophys. Res.* **86**, 3695–3707 (1981).
23. Crough, S. T. *Geophys. J. R. astr. Soc.* **55**, 451–469 (1978).
24. Sandwell, D. T. thesis, Univ. Calif. Los Angeles (1981).
25. MacDonald, G. A. *Mem. geol. Soc. Am.* **116**, 477–522 (1978).
26. Jackson, E. D., Silver, E. A., & Dalrymple, G. B. *Bull. geol. Soc. Am.* **83**, 601–617 (1972).
27. Chen, C. Y. & Frey, F. A. *Nature* **302**, 785–789 (1983).
28. Sen, G. *Earth planet. Sci. Lett.* **62**, 215–228 (1983).
29. Estill, R. E. thesis, Univ. Hawaii (1979).
30. Sharp, A. D. L., Davis, P. M. & Gray, F. *Nature* **287**, 587–591 (1980).
31. Zucca, J. J., Hill, D. P. & Kovach, R. L., *Bull. seism. Soc. Am.* **72**, 1535–1550 (1982).
32. Ellsworth, W. L. & Koyanagi, R. Y. *J. geophys. Res.* **82**, 5379–5394 (1977).
33. Sparks, R. S. J., Meyer, P. & Sigurdsson, H. *Earth planet. Sci. Lett.* **46**, 419–430 (1980).
34. Chase, T. E., Menard, M. W. & Mammerrickx, J. *Bathymetry N. Pac. Charts* **7**, 8 (Scripps Institution of Oceanography, San Diego, 1970).
35. Watts, A. B. & Talwani, M. *Geol. Soc. Am. Spec. Map Ser.* MC-9 (1975).

Novel type of crust produced during continental rifting

A. Nicolas

Laboratoire de Tectonophysique, 2 rue de la Houssinière, 44072 Nantes Cedex, France

A normal oceanic crust cannot be generated after continental crust splitting of rifts filled by a thick pile of young sediments. The magma released by the rising mantle is trapped mainly in these sediments as basalts sills and plugs. The sediments are themselves metamorphosed by the heat flux from the underlying mantle. A metasedimentary crust is formed in direct contact with an abnormal mantle ($V_p \approx 7.5 \text{ km s}^{-1}$) produced by partial serpentinization.

THE transition from continental to oceanic crust in a rift may occur in a number of different ways. One model involves the stretching of the continental crust to a thickness comparable with that of the oceanic crust, followed by a transition to this new oceanic crust. This model has been documented by the study of the Atlantic Ocean passive margins^{1,2}. Another model predicts a brittle behaviour of the continental crust which would be progressively invaded by mafic dykes, its extension being that of the accumulated thickness of the dykes, and would finally yield without stretching along the mechanically weakened zone of dyke intrusion, passing more abruptly to oceanic crust^{3,4}. Coleman⁵ describes an intermediate behaviour of dykes and igneous bodies with crustal extension and multiple intrusions which seem to produce locally a complete disruption of the continental crust on the Arabian side of the Red Sea.

In these models, the nature of oceanic crust is little questioned, despite geophysical records which are not always those of typical oceanic crust. It is also implicitly accepted that an oceanic crust can be formed whatever the sedimentary filling of the rift. Is it possible to form a normal oceanic crust, just after the splitting of continental crust, in a trough filled with kilometres of wet sediments? Saunders *et al.*⁶ have shown that in the Guaymas Basin of the Gulf of California, where the sedimentation rate can exceed $1,200 \text{ m Myr}^{-1}$, the oceanic layers 1 and 2 are partially intercalated, the young sediments being intruded and metamorphosed by many doleritic and gabbroic sills. Here, I show that in such environments the situation may be even more radical with a total absence of layer 3 and a modified upper mantle lying directly below metamorphosed sediments invaded by basaltic intrusions. Thus, I define a new type of crust, the metasedimentary crust, transitional between the continental and oceanic crusts.

I will discuss geophysical evidence from two well-known rifting environments (Salton Trough at the north end of the Gulf of California, and the southern Red Sea) and from the western Pyrénées where, starting from a similar situation during the Cretaceous⁷, it is now possible to observe the deep parts of the crust and the upper mantle peridotites, from mountain building data. The study of these peridotites supports my interpretation. Finally, I suggest an answer to the problem of the crust or mantle nature of the $7.2\text{--}7.5 \text{ km s}^{-1}$ layers often found in non-typical oceanic lithosphere at depths between crust and mantle.

Salton Trough

The Imperial Valley–Salton Sea subsiding rift corresponds to the transition of the oceanic crust of the Gulf of California to the continental crust of southern California. This rift is opening in a system of large dextral strike-slip faults limited to the north-east by the San Andreas fault. It is rapidly subsiding with the Pleistocene Colorado River delta formations now at depths of 3.7 km below the Salton Sea and 4.8 km below the Mexican border⁸.

In the trough below sedimentary layers, Fuis *et al.*⁸ define a basement ascribed to deeper sediments metamorphosed in the greenschist facies (Fig. 1). This is based on the geophysical properties of this zone and on thermal measurements in deep wells, the extrapolation of which fix the top of the basement at 300°C for a depth of 6 km. Basaltic intrusions occur in this basement. At a depth of $\sim 12 \text{ km}$, one passes into the sub-basement through a 1-km thick transition zone. This sub-basement is ascribed to a mafic layer (gabbros and/or amphibolites) representing the oceanic crust on behalf of a P-wave velocity $V_p = 7.2 \text{ km s}^{-1}$ and a modelled density of 3.1 (Fig. 1). The 7.2 km s^{-1} velocity, however, is a lower limit; velocities of $7.5\text{--}8.0 \text{ km s}^{-1}$ have also been reported. A recalculation of the same data gives an average velocity of 7.5 km s^{-1} . Here I reinterpret the sub-basement as anomalous mantle rather than mafic crust.

Fitting a P-wave velocity of 7.5 km s^{-1} and a density of 3.1 with an oceanic lithosphere formation can be done by a comparison with experimental data on ophiolite specimens or from a direct comparison with *in situ* geophysical data in the oceanic lithosphere. Ascribing to the upper sub-basement pressure-temperature ($P\text{--}T$) conditions of $\sim 4\text{--}5 \text{ kbar}$ and $400\text{--}600^\circ\text{C}$, the appropriate comparison should be done with a young crust rather than with a nascent or a mature oceanic lithosphere.

Measurements on gabbros and metagabbros at 4-kbar confining pressure and room temperature give a wide range of velocities with mean values respectively of 7.14 km s^{-1} and 6.92 km s^{-1} for a few west American ophiolites⁹; data from Oman¹⁰ or from the Bay of Islands^{11,12} do not differ significantly. Correction for an assumed temperature of 500°C would lower these values by $\sim 5\%$ (ref. 13). Density measurements on gabbros and metagabbros at normal $P\text{--}T$ conditions on the same specimens as above give 2.95 and 2.92, respectively. Data from the

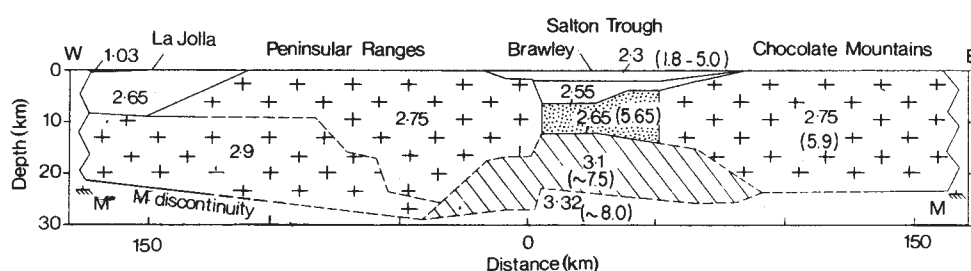


Fig. 1 Cross-sections through the Salton Trough integrating seismic (velocities in bracketed numbers) and gravimetry (plain numbers) data. Dots, metasedimentary basement; crosses, continental basement; oblique hatches, sub-basement (after Fuis *et al.*⁸). The 7.5 km s^{-1} velocity adopted for the sub-basement is derived from ref. 8 and from recalculated velocities using their data.

Table 1 P-wave velocities in layer 3 and upper mantle of young lithospheres

	Reykjanes ³⁸	Iceland ³⁹	Galapagos ⁴⁰	EPR* 10° S ⁴⁰	EPR Siqueiros ⁴¹	EPR 12° N ^{42,43}	EPR 14° N ⁴⁴	Gulf of Aden ⁴⁵	Kenya ⁴⁶
V_p in layer 3 km s ⁻¹	6.5	6.5	6.5	~6.8	7.0	6.8–7.2	6.8–7.0	6.3	
V_p in upper mantle km s ⁻¹	7.3–7.4	7.2	7.5	7.5	7.8	8.0	7.75	7.2	7.3–7.5

* EPR, East Pacific Rise.

Table 2 Comparisons between Salton Trough P-wave velocities and other estimates

	Salton Trough sub-basement 4 kbar/500 °C	Gabbro metagabbro 4 kbar/500 °C	Layer 3 young crust	14% Serpentinized peridotite 4 kbar/500 °C	Upper mantle young lithosphere
V_p km s ⁻¹	~7.5	6.57–6.78	6.3–7.2 (6.7)	7.34	7.4–8.0 (7.5)
D g cm ⁻³	3.1	2.85–2.88 ≤ 3.08		3.12	

Figures in parentheses represent average values from Table 1.

Bay of Islands and Oman show comparable densities with lower gabbros reaching densities of 3.15. Recalculating these densities for $P = 4$ kbar and $T = 500$ °C with compressibility and thermal expansion coefficients of 1×10^{-6} bar⁻¹ and 24×10^{-6} deg⁻¹, respectively, would lower these densities by ~0.07, bringing them in the range of 2.88 and 2.85, with the maximum density lowered to 3.08.

An ultrasonic compressional wave measured along the main structural directions on a peridotite with 14% serpentine at 4 kbar gives an averaged velocity of 7.63 km s⁻¹ (ref. 14). The results are similar using Christensen's¹⁵ analytical data: 7.6 km s⁻¹ for a peridotite (2/3 olivine, 1/3 enstatite) with 15% serpentine. Using the temperature derivative calculated for a comparable peridotite¹⁶, one finds for $P = 4$ kbar and $T = 500$ °C, a velocity of 7.35 km s⁻¹. The density measured on the same specimen was 3.2, which after the pressure and temperature correction, becomes 3.12.

P-wave velocity data for young oceanic lithosphere and rifts have been compiled in Table 1. When possible, the data for age zero have been avoided. Velocities in the upper mantle are commonly anisotropic, in which case they have been averaged.

Table 1 indicates that anomalous mantle velocities ~7.5 km s⁻¹ are common below young oceanic crust where they are ascribed to the strong temperature effect on this velocity. With ageing, these velocities would increase to the more common 8.2 km s⁻¹ value¹⁷. Table 2 shows that a peridotite with 10–20% serpentine fits equally well the data for Salton Trough.

Considering the fact that in the Salton Trough a source of water is found in the overlying sediments undergoing metamorphic reactions, I suggest that the upper part of the sub-basement and the transition zone are composed of partly serpentinized peridotites. The serpentinization decreases downward partly because it is difficult for the water to penetrate deep into the peridotites, but principally because the upper temperature limit for serpentinization (500 °C) is met at shallow depth in the sub-basement. At higher temperatures, tremolite-chlorite and talc can be present, contributing to the thermal effect and maintaining low velocities.

Southern Red Sea

Although the geophysical data are older and less tractable, a comparable situation with the Salton Trough seems to exist in the southern Red Sea, south of the 17° parallel. The magnetic anomaly pattern suggests the existence of oceanic crust in the central and southern Red Sea^{18,19}. This is confirmed for the central Red Sea by seismic refraction data²⁰. There, material with a mean seismic velocity of 4.3 km s⁻¹ and thought to be evaporites, overlies a layer of mean velocity 6.6 km s⁻¹, thought to be oceanic crust.

In the southern Red Sea, seismic refraction and gravimetry profiles²¹ and structural reconstitutions²² in the 50–100-km-wide central trough define sedimentary layers 2–5 km in thickness with velocities from 2.75 to 4.48 km s⁻¹, overlying directly a 7.16–7.31 km s⁻¹ domain. These overlying layers are evaporite sediments²² associated with mafic formations. Seismic reflection profiles²³ suggest the occurrence of volcanic strata and plugs in the sediments. There is a general consensus that the underlying domain is mantle^{22–24} whose anomalously low velocities are related to the locally high thermal gradient. Similarly, low mantle velocities have been reported from the adjacent areas²⁴ (Afar, Danakil, Jibuti, Gulfs of Tajura and of Aden).

The abundance of mafic intrusions in the southern Red Sea and along its margins can account for the magnetic anomalies mentioned above and do not necessarily imply the presence of an oceanic crust that is not supported by the available geophysical record. Equating the 7.2 km s⁻¹ basement with anomalous mantle, we conclude that this mantle is directly overlain by a sedimentary cover with mafic sills or flows and plugs. New seismic refraction profiles would be necessary to test this interpretation.

Western Pyrénées

During the Pyrenean orogeny, small and deep sedimentary basins, possibly analogous to those described above⁷ have been squeezed, making it possible to obtain some information on the nature of their substratum.

These sedimentary troughs were filled by flysch during their opening. Thus, 3–6-km-thick flysch was accumulated rapidly in the western Pyrénées where the troughs are best developed. After a period of Triassic continental rifting followed by a relative quiescence, the rifting was reactivated during the Upper Jurassic–Lower Cretaceous²⁵, with the formation of the troughs under consideration. The precise mechanism for their opening is still unknown. The troughs may open by a pull-apart mechanism related to the essentially sinistral motion of the Iberia block with respect to Europe⁷, or there may be a period of extensional tectonics within Iberia, followed by the strike-slip motion and the compression responsible for the mountain building event²⁶. The first model seems to better explain the larger opening observed in the western Pyrénées. The content of the troughs has been metamorphosed. Temperatures as high as 600 °C were attained locally during the Albian–Cenomanian phase of the rifting²⁷. During the subsequent compressional phase, they were still high enough to produce metamorphic effects.

In the eastern Pyrénées, the continental crust was probably only attenuated during stretching²⁸. The mainly calcareous sediments from the basins contain small mafic and lherzolitic bodies, the latter associated with granulite slivers derived from lower

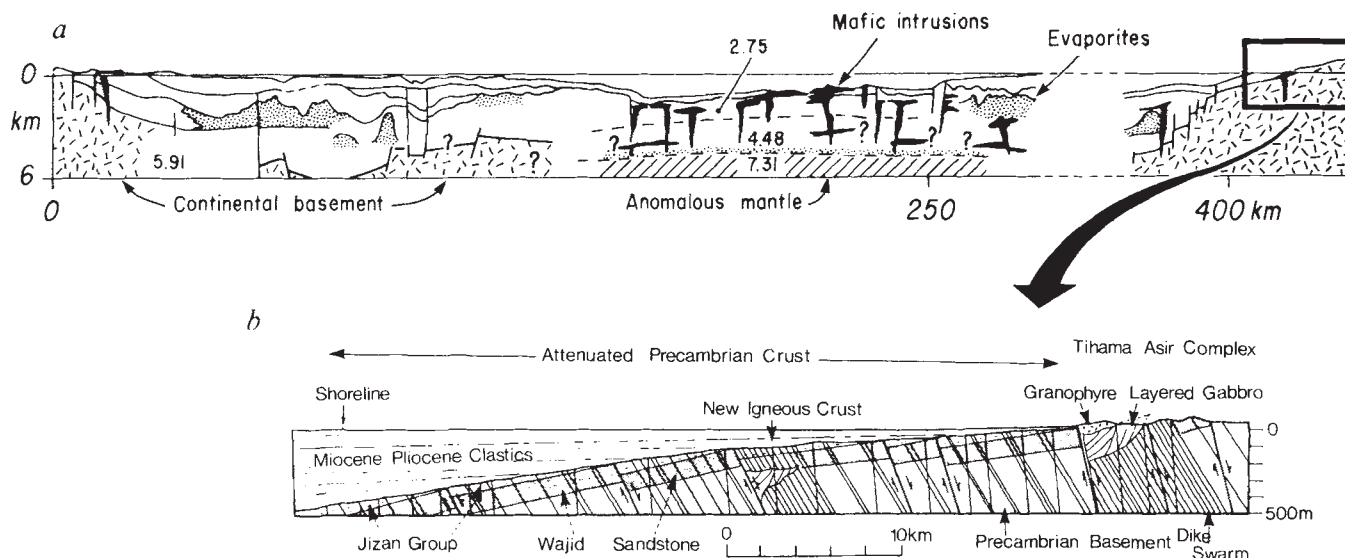


Fig. 2 Cross-section through *a*, southern Red Sea (near 17° parallel) and *b*, Saudi Arabia passive margin, illustrating the contrast in mafic magmatic activity: in the Red Sea trough, mafic sills, plugs and flows and along the passive margin, local creation of oceanic crust (ophiolites). *a*, Modified from Mulder *et al.*²³ by adding main seismic refraction results²¹ and presumed deeper mafic intrusions in the 4.48 km s⁻¹ layer; *b*, reproduced from Coleman⁵.

continental crust and modified at depth by the thermal effects of the lherzolites²⁹. These numerous lherzolite bodies are remarkably fresh with a plastic deformation indicative of large strain in the $T \approx 1,000^\circ\text{C}$ range, but locally down to $650\text{--}700^\circ\text{C}$ (ref. 30), which we ascribe to their ascent.

In the western Pyrénées, the continental crust was probably disrupted^{26–28}. This is suggested by the nature of the troughs and the large gravity anomaly. The small peridotite massifs upthrust during the orogeny are more harzburgitic (F. Boudier and F. Cordellier, personal communication). They are also largely serpentized (lizardite), essentially at low temperatures and in static conditions, but with relicts of a high temperatures earlier stage in the greenschist facies (antigorite-tremolite-chlorite). The mafic intrusions are mainly ophites which are metamorphosed tholeiitic dolerites, intrusive within the Triassic evaporites in remarkably abundant and large (kilometre-sized) sills and plugs, and alkali-basalt sills and plugs related to the Albian rifting. The ophites are considered to be related to the Triassic period of rifting³¹ because they do not intrude younger formations. But this may result from the location of Triassic evaporites at the bottom of the troughs or to their mechanical properties; the ophites could be related to the Cretaceous rifting as suggested by other authors³².

Discussion

A new analysis of the geophysical data on the Salton Trough and southern Red Sea rifts leads to the conclusion that after splitting of the continental crust a new lithosphere can be created with a metasedimentary section directly above the mantle formations and without a normal oceanic crust. The existence of this metamorphism is deduced from heat-flow data⁸. It is also imprinted in the Red Sea evaporitic sediments in contact with the Zabargad Island intrusion of hot spinel-feldspar lherzolites³³. The metasediments are intruded by mafic sills and plugs²³. In the Red Sea, there is a striking contrast between the style of mafic magmatism in the continental crust (Fig. 2*b*) and in the evaporitic sediments of the trough (Fig. 2*a*).

The Mesozoic Pyrenean troughs seem to represent a similar situation, at least in the western part where the continental disruption was probably complete, a stage not attained in the eastern part. This interpretation is supported by the comparative study of the intruded peridotite bodies.

The scenario proposed for the western Alps lherzolites³⁴ can be applied to spinel lherzolites from the eastern Pyrénées, which are very similar. They would be derived from a diapiric asthenosphere uprise below a continental rift. Because of a particularly

slow ascent rate in relation to a presumed slow ($<1\text{ cm yr}^{-1}$) opening rate, the partially molten asthenosphere meets the thermal conduction front of the lithosphere during its ascent at $\approx 30\text{ km}$, where the melts carried so far start crystallizing³⁴. This explains the comparatively fertile character of the lherzolites, their equilibration in the spinel-lherzolite field, the limited amount of known associated magmatism and the low temperature characteristics of their subsequent plastic deformation, which records their further ascent and possibly the first stages of the rift closure. It has been proposed recently³⁵ that this lherzolite diapir has disrupted the continental crust and intruded through the rift sediments. This interpretation, proposed in the case of feldspar-lherzolites^{34,36}, seems unlikely for the spinel lherzolites considered here.

In the western Pyrénées, the occurrence of harzburgitic lherzolites is symptomatic of the creation of oceanic lithosphere³⁴ and of substantially larger spreading rates there than in the east³⁶. In Boillot's model²⁶, it is possible to estimate the opening rate at $\approx 1.3\text{ cm yr}^{-1}$, which is sufficient to generate an oceanic lithosphere in terms of the amount of melting (20%) and of the harzburgitic nature of the residual mantle³⁶. However, the oceanic mantle would be directly overlain by metamorphosed sediments, as suggested by the absence of oceanic crust fragments associated with the harzburgite bodies. The large and numerous ophite intrusions could represent the mafic products generated by the mantle melting, but this would imply that they are not Triassic (see above). The extensive serpentization of the harzburgite bodies, contrasting with its scarcity in eastern lherzolites, would be partly achieved at high temperatures by water contamination from the overlying metasediments and partly during and after the squeezing up of the peridotites through the trough's formation. This is recorded by the serpentization sequence, showing first a high temperature alteration (antigorite-tremolite-chlorite), followed by a low temperature one (lizardite).

Such a partial and early serpentization of the mantle section beneath metamorphosed sediments may explain the seismic and gravimetric data obtained in the Salton Trough⁸. In particular, the upper sub-basement could correspond there to a peridotite with 15–20% serpentine.

Conclusions

Based on the data presented here, I propose that in certain circumstances a new type of crust can be created during the rifting of a continent. During the disruption of the continental crust, this new crust would be composed of a thick sedimentary

sequence, metamorphosed at its base in the greenschist or the amphibolite facies depending on its thickness and be invaded by sills and plugs of kilometre-sized mafic intrusions. The immediately underlying anomalous mantle ($V_p \approx 7.5 \text{ km s}^{-1}$) would be composed of hot peridotites and/or of partially serpentinized peridotites. The latter interpretation agrees with the one proposed by Hess³⁷, in which a lower oceanic crust is formed by serpentinized peridotites. The formation of this special crust would require conditions in which the creation of a normal oceanic crust would be impossible. Such conditions could occur when the transition from continental to oceanic lithosphere occurs in a trough filled with a thick pile of young sediments, a situation opposed to that of a starved rift¹. The basaltic magma released by the underlying mantle diapir into the immature sediments does not meet conditions to create a volcanic cap below which a plutonic section can fractionate. On the contrary, as already proposed by Saunders *et al.*⁶, the magma reacts with these sediments and is frozen in them as sills and plugs, thus

contributing to the metamorphism of the sediment. Dolerite and gabbro sills material and metamorphosed sediments have been recovered by DSDP drilling in the Gaymas Basin of the Gulf of California⁶ which, as mentioned above, seems to correspond to an appropriate environment.

It may be of interest to contrast the situation on the Arabian edge of the Red Sea and in the trough itself (Fig. 2). On the edge, the disruption of the continental crust with a little and old sedimentary cover leads directly to the creation of typically ophiolitic segments of oceanic crust, whereas in the trough where young and wet sediments have accumulated, the geophysical record does not show the existence of such oceanic crust but suggests a direct contact between sediments and serpentinized mantle.

This article stemmed from discussions with M. Mattauer. I thank X. Le Pichon and J. A. Karson for helpful reviews. Support from Centre National de la Recherche Scientifique (ERA 547) is acknowledged.

Received 23 October 1984; accepted 14 February 1985.

- Montadert, L., Roberts, D. G., De Charpal, O. & Guennoc, P. *Init. Rep. DSDP* **48**, 1025-1060 (1979).
- Talwani, M. & Langseth, M. *Science* **213**, 22-31 (1981).
- Turcotte, D. L., Ahern, J. L. & Bird, J. M. *Tectonophysics* **42**, 1-28 (1977).
- Mohr, P. *Tectonophysics* **94**, 509-528 (1983).
- Coleman, R. G. *Proc. 27th Int. geol. Congr., Moscow* **9**, 221-239 (1984).
- Saunders, A. D., Fornari, D. J. & Morrisson, M. A. *J. geol. Soc.* **139**, 335-346 (1982).
- Choukroune, P. & Mattauer, M. *Bull. Soc. géol. Fr.* **20**, 689-700 (1978).
- Fuis, G. S., Mooney, W. D., Healy, J. H., McMechan, G. A. & Luther, W. J. *J. geophys. Res.* **89**, 1165-1189 (1984).
- Christensen, N. I. *Tectonophysics* **47**, 131-157 (1978).
- Christensen, N. I. & Smewing, J. D. *J. geophys. Res.* **86**, 2545-2556 (1981).
- Salisbury, M. H. & Christensen, N. I. *J. geophys. Res.* **83**, 805-817 (1978).
- Karson, J. A., Collins, J. A. & Casey, J. F. *J. geophys. Res.* **89**, 6126-6138 (1984).
- Christensen, N. I. in *Handbook of Physical Properties of Rocks* Vol. 2 (ed. Carmichael, R. S.) 1-227 (CRC, Florida, 1982).
- Peselnick, L., Nicolas, A. & Stevenson, P. R. *J. geophys. Res.* **79**, 1175-1182 (1974).
- Christensen, N. I. *J. geophys. Res.* **71**, 5921-5931 (1966).
- Peselnick, L. & Nicolas, A. *J. geophys. Res.* **83**, 1227-1235 (1978).
- Bibee, L. D. & Shor, G. G. Jr *Geophys. Res. Lett.* **11**, 639-642 (1976).
- Roeser, H. A. *Geol. Jb D13*, 131-153 (1975).
- Girdler, R. W. & Styles, P. in *Afar Depression of Ethiopia* (eds Pilger, A. & Rosler, A.) 138-154 (Schweizerbart, Stuttgart, 1976).
- Tramontini, C. & Davies, D. J. *R. astr. Soc.* **17**, 225-241 (1969).
- Drake, C. B. & Girdler, R. W. *J. R. astr. Soc.* **8**, 473-495 (1964).

- Lowell, J. D. & Genik, G. J. *Am. Ass. petr. Geol. Bull.* **56**, 247-259 (1972).
- Mulder, C. J., Lehner, P. & Allen, D. C. K. *Geologie Mijnb.* **54**, 208-221 (1975).
- Mohr, P. A. *Rev. Earth planet. Sci.* **6**, 145-172 (1978).
- Brunet, M. F. *Geol. Mag.* **121**, 421-428 (1984).
- Boillot, G. *Geol. Mag.* **121**, 407-412 (1984).
- Albarède, F. & Michard-Vitrac, A. *Earth planet. Sci. Lett.* **40**, 327-332 (1978).
- Mattauer, M. *C. r. heb. Séanc. Acad. Sci., Paris* **300**, 71-74 (1985).
- Monchoux, P. *Contrib. miner. Petrol.* **37**, 47-64 (1972).
- Conquére, F. & Fabries, J. in *Kimberlite II: the Mantle and Crust-Mantle Relationships* (ed. Kornprobst, J.) 319-331 (Elsevier, Amsterdam, 1984).
- Azambre, B., Rossy, M. & Elloy, R. *Bull. Soc. géol. Fr.* **23**, 263-269 (1981).
- De Sitter, L. U. & Zwart, H. J. *Leid. geol. Meded.* **26**, 1-49 (1961).
- Kurat, G., Niedermayr, G., Prinz, M. & Brandstater, F. *Terra Cognita* **2**, 240 (1982).
- Nicolas, A. in *Kimberlite II: the Mantle and Crust-Mantle Relationships* (ed. Kornprobst, J.) 333-345 (Elsevier, Amsterdam, 1984).
- Vielzeuf, D. & Kornprobst, J. *Earth planet. Sci. Lett.* **47**, 87-96 (1984).
- Boudier, F. & Nicolas, A. *Earth planet. Sci. Lett.* (submitted).
- Hess, H. H. in *Petrologic studies* 599-620 (Geol. Soc. Am., 1962).
- Talwani, M., Windisch, C. C. & Langseth, M. G. Jr *J. geophys. Res.* **76**, 473-517 (1971).
- Palmasen, G. & Saemundsson, K. A. *Rev. Earth planet. Sci.* **2**, 25-50 (1974).
- Bibee, L. D. thesis (Univ. Calif., San Diego (1979)).
- Rosendahl, B. R. & Dorman, L. M. *Init. Rep. DSDP* **54**, 23-36 (1980).
- Lewis, B. T. R. & Garmany, J. D. *J. geophys. Res.* **87**, 8417-8425 (1982).
- Bratt, S. R. & Purdy, G. M. *J. geophys. Res.* **89**, 6111-6125 (1984).
- Lewis, B. T. R. & Snyderman, W. E. *Tectonophysics* **55**, 87-105 (1979).
- Laughton, A. S., Whitmarsh, R. B. & Jones, M. T. *Phil. Trans. R. Soc. A267*, 227-266 (1970).
- Long, R. E. & Backhouse, R. W. *Geophys. J. J.* **44**, 677-688 (1976).

Heritable formation of pancreatic β -cell tumours in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes

Douglas Hanahan

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

Following the transfer into fertilized mouse eggs of recombinant genes composed of the upstream region of the rat insulin II gene linked to sequences coding for the large-T antigen of simian virus 40, large-T antigen is detected exclusively in the β -cells of the endocrine pancreas of transgenic mice. The α - and δ -cells normally found in the islets of Langerhans are rare and disordered. Well-vascularized β -cell tumours arise in mice harbouring and inheriting these hybrid oncogenes.

INSULIN, a polypeptide hormone involved in the control of carbohydrate metabolism, is synthesized and stored in the β -cells of the endocrine pancreas, from which it is released into the bloodstream in response to various signals. The endocrine pancreas is comprised of isolated islands or islets, which are dispersed throughout the larger mass of the exocrine pancreas; the islets consist of four major cell types, α , β , δ and PP, which synthesize the hormones glucagon, insulin, somatostatin and pancreatic polypeptide, respectively^{1,2}. Rat and human insulin genes have been cloned and analysed³⁻¹³. There are two non-

allelic rat insulin genes that differ in nucleotide sequence and in the number of introns; the rat insulin I and II genes have one and two introns, respectively^{5,6}. The two genes share a region of considerable sequence homology in the 5' flanking region, extending ~300 base pairs (bp) upstream of the point of transcriptional initiation¹²; the human gene is 75% homologous over the first 240 bp of this region⁸. Expression of hybrid genes using the 5' flanking regions of both the human insulin and rat insulin I genes occurs when they are introduced transiently into cultured cells derived from an insulinoma¹¹.

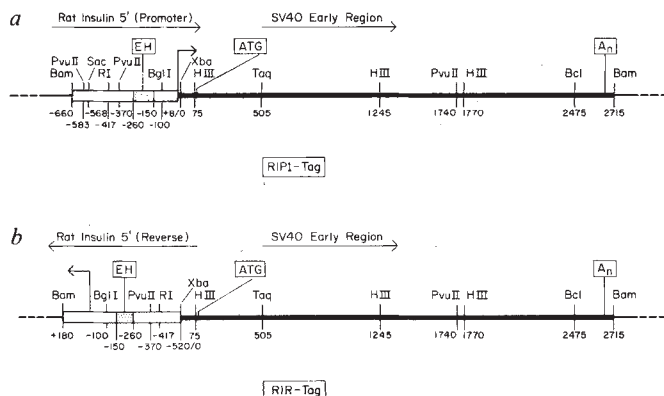
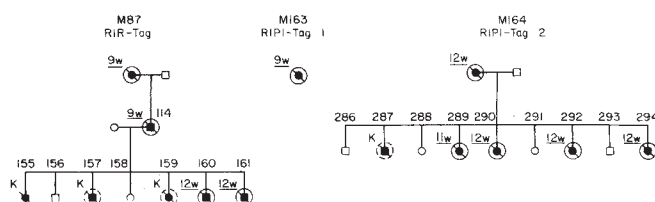
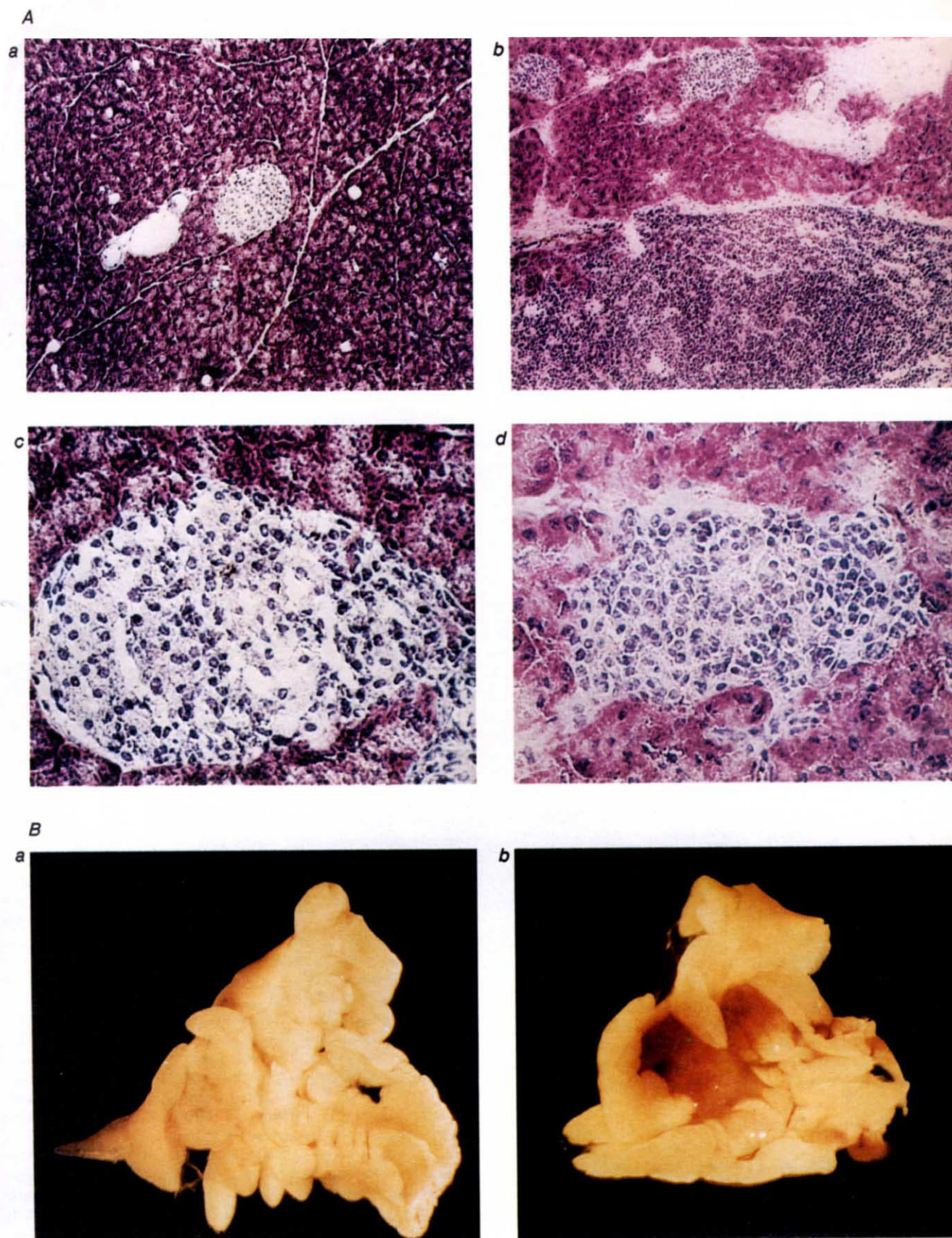


Fig. 1 Structure of recombinant insulin/large-T antigen genes. The plasmids pRIP1-Tag (a) and pRIR-Tag (b) consist of the coding information of the SV40 early region fused to sequences derived from the 5' flanking region of the rat insulin II gene. Boxed regions denote insulin gene flanking DNA and the thick solid line refers to the SV40 early region. The cap site where insulin gene transcription initiates is indicated by an L-shaped arrow, and the associated transcriptional enhancer element¹¹ (EH) is shown (the insulin gene promoter, which lies between them, is not indicated). The points of SV40 large-T antigen translation initiation and transcription termination are indicated by boxed ATG and A_n respectively. The BglII site in SV40 was converted to an XbaI site using an oligonucleotide linker, and the Xba/BglII to Bam fragment comprising the SV40 early region²⁴ was linked to two different orientations of the insulin gene flanking region. This fragment of the SV40 early region includes the early mRNA cap sites, but lacks all known components of the SV40 early promoter. For RIP1-Tag, an Xba linker was inserted into the DdeI site at +8 bp relative to the point of initiation of insulin gene transcription, and the Bam to Dde/Xba fragment from -660 to +8 was combined with the Xba/Bam fragment of the SV40 early region and inserted as a Bam linear into a derivative of pBR322 that lacks the R1 site, and includes a lacUV5 promoter in place of the R1 to Bam fragment of pBR. For the plasmid pRIR-Tag, which carries the insulin flanking region in the opposite orientation, a Dde site at -520 was converted to an Xba site, and the Bam/Xba fragment extending from +180 nucleotides past the cap to -520 bp upstream was combined with the structural gene for large-T antigen and inserted into the same plasmid vector. The two hybrid genes were similarly prepared for microinjection, except that for RIR-Tag, the Bam insert was first purified from the plasmid vector by gel electrophoresis and electroelution, followed by passage over a DEAE-Sephacel column and RIP1-Tag was linearized using SalI. The DNAs were extracted with phenol, chloroform, precipitated in 70% ethanol, resuspended in 10 mM NaCl, 5 mM Tris pH 7.4, 0.1 mM EDTA and passed through a 1-ml spin column of Sepharose CL6B, after which the DNA was diluted to concentrations of ~75 copies per pl for pRIP1-Tag/Sal or ~20 copies per pl for the RIR-Tag/Bam insert. The DNAs were injected into fertilized one-cell embryos essentially as described^{14,16}, except that the injected embryos were cultured overnight to the two-cell stage before being re-implanted into the oviducts of pseudopregnant females. The F₂ embryos were derived from matings of B6D2F1 (C57BL/6J × DBA/2J) males and females, obtained from the Jackson laboratory.

Here I seek to address aspects of insulin gene expression by the introduction of recombinant insulin genes into the mouse germ line, using microinjection of DNA into fertilized mouse embryos. This technique has been used previously to examine the control and consequences of expression of a number of genes¹⁴⁻²³, including two oncogenes. The complete simian virus 40 (SV40) early region, including the viral transcriptional enhancer and promoter, is not expressed at detectable levels in normal tissues of transgenic mice harbouring it, but choroid plexus tumours heritably arise, and these tumours produce high levels of large-T antigen¹⁹. The cellular proto-oncogene *c-myc*, when linked to the long terminal repeat of mouse mammary tumour virus, is expressed in a number of tissues, including mammary cells, and two lines of mice heritably develop breast tumours²⁰.

The recombinant oncogenes used here consist of regulatory



**Fig. 3**

Hybrid gene constructs

Two hybrid genes were constructed to combine DNA 5' to the coding sequences of rat insulin II gene with the large-T antigen structural information. This 5' flanking region includes the insulin gene promoter as well as a transcriptional enhancer element that stimulates initiation of transcription in cultured insulinoma cells but not in fibroblasts¹¹. The first configuration, denoted *RIP1* (rat insulin promoter 1) carries the insulin control region aligned to promote transcription of the large-T antigen gene (Fig. 1). The two genes were linked together in regions corresponding to the 5' untranslated portion of each messenger RNA. The resulting hybrid gene, *RIP1-Tag*, thus comprises a complete transcription unit with 660 bp of sequence located 5' to the point of transcriptional initiation of the insulin gene linked to protein coding and termination information from the early region of SV40.

The second configuration, denoted *RIR-Tag* (rat insulin reverse), includes 520 bp of DNA 5' to the insulin gene and the first 180 bp of the transcribed portion of the gene. In this hybrid, the promoter element for insulin gene transcription is inverted with respect to the coding information for SV40 large-T antigen, and, therefore, does not comprise a fusion in the 5' untranslated regions of the two genes, but rather includes an inverted promoter and enhancer region linked to the *Tag* structural gene, as shown in Fig. 1. Both hybrid genes lack the SV40 early promoter and transcriptional enhancer element, and each includes protein coding information for SV40 small-T antigen, as well as for large-T, because the two proteins are derived from the primary transcript of the SV40 early region by differential splicing of overlapping introns²⁴.

The two fusion genes were injected into fertilized one-cell mouse embryos, which were then inserted into the oviducts of pseudopregnant female mice and allowed to develop. Four transgenic mice were produced from the *RIP1-Tag* injections and one from the *RIR-Tag* injections. The apparent heterozygous copy number (per diploid genome) of the acquired DNA ranged from ~1 for mouse 87, which carries *RIR-Tag*, to ~5 for mouse 163, which carries and is genotyped as *RIP1-Tag1*, and 5 for mouse 164 (*RIP1-Tag2*). Two of the founder mice (165 and 168) were mosaic for the acquired transgene, using the criteria of infrequent transmission to progeny and lower apparent copy number in the founder than in its progeny. The non-mosaic progeny of M165 (*RIP1-Tag3*) have copy numbers of ~25, whereas those of M168 (*RIP1-Tag4*) have ~12 copies per diploid genome. In each case of multiple copies, genetic and biochemical analysis indicates that the genes are heterozygous and integrated in a single location (data not shown).

Phenotype of transgenic mice

During the initial breeding analysis of the insulin large-T antigen transgenic mice, a phenotype of premature death emerged. The three original non-mosaic mice died at 9–12 weeks of age. Two of these transmitted the transgene to offspring before dying, whereas the third (M163) did not. Surprisingly, this phenotype appeared with the *RIR-Tag* gene, which has the insulin promoter inverted with respect to the large-T antigen coding region, as well as with the *RIP1-Tag* gene, in which the insulin promoter is aligned to transcribe the large-T antigen gene.

The phenotype of premature death proved to be heritable and co-segregated with the acquired hybrid gene; the partial pedigree in Fig. 2 illustrates this observation. The penetrance is virtually complete in the early generations. The analysis that follows deals primarily with the progeny of the two original non-mosaic mice that transmitted—M164 (*RIP1-Tag2*) and M87 (*RIR-Tag*). Progeny of each were killed and subjected to pathological and biochemical analysis. The only obvious abnormal pathology consisted of hyperplasia of the islets of Langerhans, determined following standard histochemical staining of tissue sections. Solid tumours were observed in a few of these mice. Figure 3a shows representative pancreas sections of a normal mouse and of a transgenic mouse (M384, *RIP1-Tag2*) that carried visible

tumours; both were ~12 weeks old. Sections were stained with cresol violet, which distinguishes the endocrine islets from the surrounding exocrine tissue. Sections from M384 showed normal-sized islets, enlarged (hyperplastic) islets and solid tumours. All stained with similar characteristics, and were identifiable as islet-like regions by comparison with normal islets. Note, however, that all islet-like areas in the transgenic mouse appear more densely packed with cells than do the normal mouse islets, demonstrated by comparison of Fig. 3A c and d.

Analysis of the mice listed in Fig. 2 shows that all transgenic mice in these two lineages have the phenotype of islet hyperplasia (if killed) and/or premature death. Before their sudden death, the mice seem normal in appearance and activity. The pedigree stops at the generation during which the mice were placed on a high sugar diet to counteract indications that they were hypoglycaemic, which is a probable consequence of islet hyperplasia, and a possible cause of sudden premature death. Subsequent generations, maintained on this diet, now live somewhat longer and heritably develop solid tumours of the pancreas between 10 and 20 weeks of age. An example of a pancreas containing several tumours is shown in Fig. 3B. These tumours eventually comprise a significant fraction of the mass of the pancreas and are highly vascularized. Individual tumours can reach a size of 5 mm in diameter, with up to five or six separate tumours visible in a pancreas. There is no indication that the tumours metastasize other organs.

Cell-type specificity

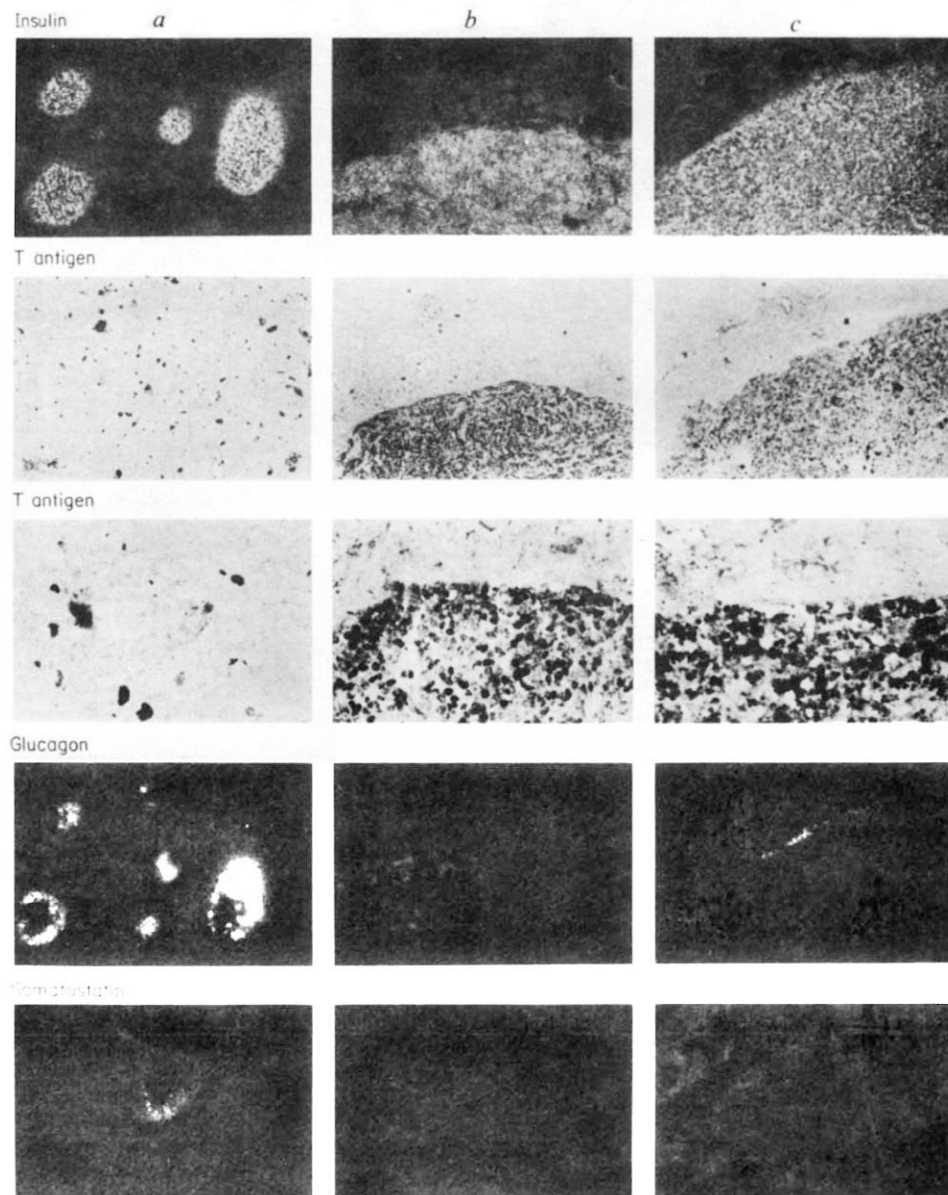
To evaluate the observed hyperplasia of the islets of Langerhans, I examined normal and transgenic mouse pancreas for expression of both SV40 large-T antigen and polypeptide hormones characteristic for the major cell types. Thin sections were prepared from fresh frozen pancreas obtained from two transgenic mice (M384, *RIP1-Tag2*, and M306, *RIR-Tag*) carrying visible pancreatic tumours as well as from a normal mouse. The sections were analysed by immunostaining using antisera to glucagon, insulin and somatostatin and each visualized with rhodamine-conjugated second antibody to identify the α -, β - and δ -cell types of the islets. Large-T antigen was identified using both a polyclonal antiserum and monoclonal antibodies, each visualized by horseradish peroxidase-conjugated second antibody. Figure 4a shows representative immunostains of normal (non-transgenic) islets for the presence of insulin, large-T antigen, glucagon and somatostatin. The islets are comprised predominantly of α - and β -cells with fewer δ -cells; no cells react with antibodies to large-T antigen. Controls using second antibodies alone show no specific staining (not shown).

The islet tumours (Fig. 4b, c) in contrast, seem to consist of solely of insulin-producing cells, identifiable as β -cells by immuno- and histochemical staining. Glucagon- and somatostatin-producing cells are rare if not completely absent from the tumours. Characteristic localization of large-T antigen to the nucleus is observed in the islet tumour cells; it is clearly absent from adjacent exocrine cells. Evidence from both polyclonal serum (shown) and monoclonal antibodies (not shown) indicates the same pattern of large-T antigen expression. By these criteria, the islet tumours are essentially pure populations of proliferating β -cells. Occasional clusters of α -cells, perhaps remnants of an islet, can be detected on the edges of a tumour. Thus, both hybrid oncogenes specifically induce proliferation of the β -cells of the islets of Langerhans, eventually resulting in the appearance of pure β -cell tumours. In comparison, naturally occurring insulinomas are characteristically a mixed population of both β - and δ -cells, producing insulin and somatostatin, respectively^{25,26}.

Tissue-specific expression

The ability of the insulin control region to mediate tissue-specific expression of the SV40 large-T antigen following integration of the hybrid genes into the mouse germ line was assessed by

Fig. 4 Analysis of islet cell hormone and large-T antigen expression in a normal mouse and two tumour-bearing transgenic mice. Each column shows an analysis of one mouse: *a*, normal (non-transgenic); *b*, 384 from the *RIP1-Tag2* line; *c*, 306 from the *RIR-Tag* line. In each case, near (but not necessarily adjacent) thin sections were analysed for the presence of the specified antigen by using an antiserum to that antigen followed by a conjugated second antibody to visualize the first. The first row examines insulin, the second and third rows large-T antigen, the fourth row glucagon and the fifth row somatostatin. Magnification is $\times 50$ except for the third row, which is $\times 200$. The use of second antibodies alone gives no specific staining. In *a*, insulin and glucagon plates are of the same cluster of islets, whereas the somatostatin and large-T antigen are of other islets, because that cluster was not present in those sections. The immunostains for the hormones were visualized with rhodamine-conjugated second antibody and epifluorescent illumination, whereas the immunostain for large-T antigen used peroxidase-conjugated second antibody, and was viewed under bright-field illumination. *b*, *c*, Orientation of all the plates is the same. The islet tumour is on the bottom, with adjacent exocrine tissue above it. Column *b* is aligned so the tumour is a centred arc, whereas the plates in *c* have the tumour ascending to the right. The glucagon and somatostatin plates of *b* and *c* are of sections near those used for insulin and large-T antigen; the outline of the exocrine/tumour boundary can be seen in each. The $\times 200$ plates of large-T antigen also show an exocrine/tumour boundary, in which peroxidase staining occurs in nuclei in the tumour tissue but not in the exocrine tissue.

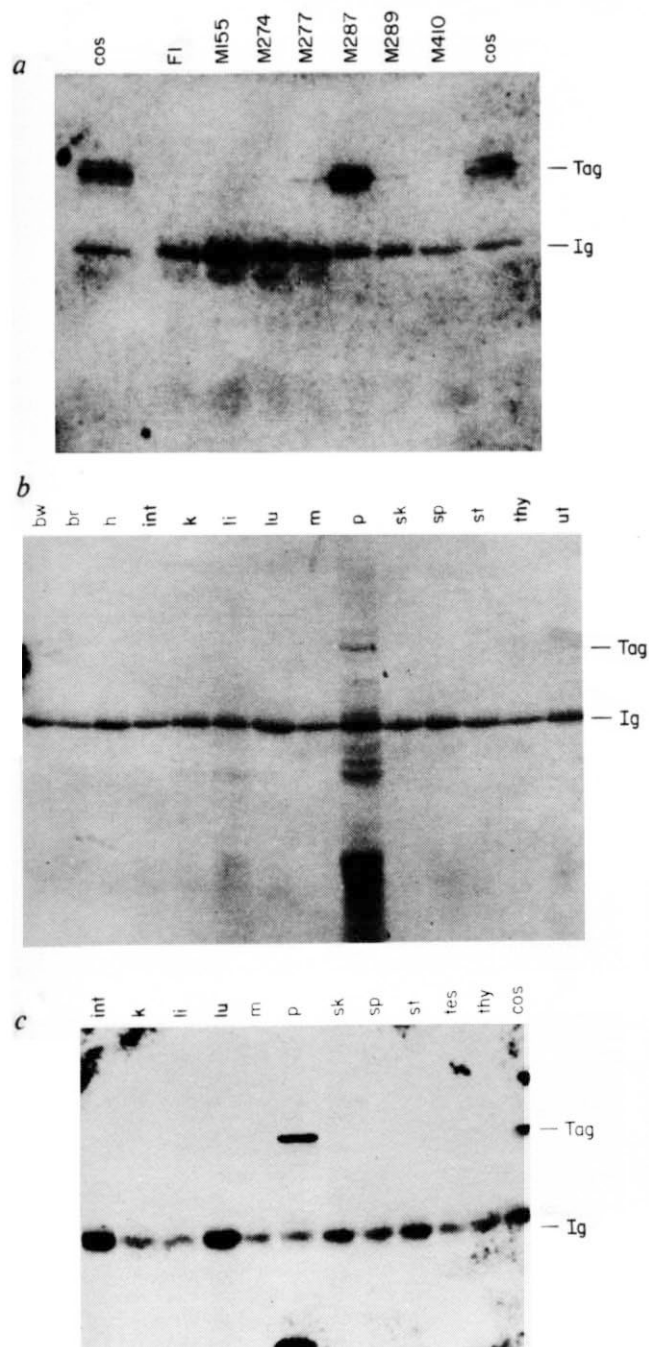


Methods. 14- μ m sections were taken

as described in Fig. 3 legend, except that for the large-T antigen immunostaining the sections were not fixed, but rather air dried on gelatinized slides. The hormone immunostains were performed on paraformaldehyde-fixed sections as in Fig. 3. Sections were rehydrated in Net-Gel (150 mM NaCl, 5 mM EDTA, 50 mM Tris pH 7.4, 0.25% gelatin, 0.02% NaN_3 , 0.05% NP40), then treated for 30 min in blocking solution (45% Net-Gel, 45% Dulbecco's modified Eagle's medium, 10% fetal bovine serum, 1% bovine serum albumin (BSA), 0.5% NP40). The sections were then incubated overnight at room temperature in the first antibody, in a binding solution (blocking solution without BSA), washed several times for 30 min in Net-Gel, incubated with second antibody in binding solution for 3 h at room temperature, and again washed in several changes Net-Gel. The sections treated with rhodamine-conjugated second antibody were taken through graded alcohol into xylene and mounted under coverslips in Gurr Fluoromount +2% diazabicyclooctane. The large-T antigen sections were visualized with horseradish peroxidase-conjugated second antibody, which was developed by treatment for 10 min in 0.25 mg ml^{-1} diaminobenzidine, 3 mg ml^{-1} nickel sulphate, 0.003% H_2O_2 . The slides were then washed several times in Net-Gel and mounted. The antisera, sources and dilutions are as follows: guinea pig anti-porcine insulin (gift of R. Santerre) 1:5,000; rabbit anti-large-T antigen (a gift of D. Lane) 1:2,000; rabbit anti-human glucagon, Dako A565, 1:1,000; rabbit anti-human somatostatin, Dako A566, 1:1,000; horseradish peroxidase-conjugated swine anti-rabbit IgG, Dako P217, 1:200; rhodamine-conjugated goat anti-guinea pig IgG, Capel 2207-0081, 1:200; rhodamine-conjugated sheep anti-rabbit IgG, Capel 2212-0084, 1:200.

protein blotting analyses, using monoclonal antibodies that specifically recognize large-T antigen. Figure 5a presents a comparison of pancreas from several mice, including one normal (non-transgenic) mouse, several insulin/large-T antigen transgenic mice, and one carrying a collagen/large-T antigen fusion gene (M410) which therefore lacks the insulin gene sequences (D.H., unpublished results). M287 and M289 are siblings in the *RIP1-Tag2* lineage (see Fig. 2); M287 had large solid tumours whereas M289 did not, which illustrates the difference in levels of expression observed in mice of the same age and genotype with and without solid tumours. M277 contains immunoprecipit-

able large-T antigen, demonstrating that the *RIP1-Tag3* lineage derived from the founder M165 also expresses the fusion gene. Large-T antigen is not detectable in a normal (F_1) pancreas or in the transgenic mouse (M410) that lacks the insulin gene regulatory region. The *RIR-Tag* transgenic mouse 155 does not show expression in the blot demonstrated here, but other analyses show a low level of large-T antigen. The *RIP1-Tag4* lineage, exemplified here by M274, does not show detectable expression of large-T antigen by protein blotting. However, pancreatic tumours occasionally arise in this line, which indicates that the *RIP1-Tag4* insertion can express.



The tissue specificity of fusion gene expression was evaluated further by comparing a number of different tissues from two mice (Fig. 5b, c), progeny of the two founders M164 and M165. Both show tissue-specific expression of SV40 large-T antigen, as only pancreas is producing immunoprecipitable protein that co-migrates with authentic large-T antigen. Some degradation of the antigen in pancreatic tissue is observed routinely; it is not clear whether this represents a normal situation in the β -cells or is simply a consequence of the protein isolation procedure. Analysis of a mouse from the *RIR-Tag* lineage also shows expression only in the pancreas, as does a similar examination of the tissues of mouse 163 (*RIP1-Tag1*) (not shown). In a normal pancreas, insulin-producing β -cells comprise only ~1-2% of the total cell mass^{1,2}. Therefore, a lower limit can be established for inappropriate expression: either <~1% of the cells in another tissue could be expressing at levels comparable with those in β -cells, or all the cells in an inappropriate tissue could be expressing at levels of <1% of those observed in the islet cells. Within these limits, the expression of both

Fig. 5 Tissue specificity of transgene expression. The presence of SV40 large-T antigen in tissues was evaluated by protein blotting^{39,40}, in which tissue homogenates were immunoprecipitated with a monoclonal antibody (PAb 419) (ref. 41) to T antigen, fractionated on a 10% SDS-polyacrylamide gel, transferred to nitrocellulose by electroblotting and visualized with a second radiolabelled monoclonal antibody (PAb 416), which recognizes a different epitope on the protein⁴¹. a, Pancreas from several mice are compared. F₁ is a non-transgenic mouse, M155 an *RIR-Tag* transgenic, M274 is *RIP1-Tag4*, M277 is *RIP1-Tag3*, M287 and M289 are from siblings from the *RIP1-Tag2* lineage, with M287 having visible tumours and M289 only hyperplastic islets, and M410 is a mouse harbouring a type I collagen/SV40 large-T antigen fusion gene. To provide a control of authentic large-T antigen, total protein from a lysate of the SV40-transformed monkey cell line cos A2 (ref. 42) was immunoprecipitated and similarly treated, except that only 5% as much protein was used compared with the amount of pancreas protein. b, c, Tissue specificity of large-T antigen expression in the *RIP1-Tag2* (c) and *RIP1-Tag3* (b) lineages. Equal concentrations of each tissue were examined: bw, body wall; br, brain; h, heart; int, intestine; k, kidney; li, liver; lu, lung; m, muscle; p, pancreas; sk, skin; st, stomach; tes, testes; thy, thymus; ut, uterus. A separate analysis indicates that no large-T antigen can be detected in body wall, brain or heart of M287 (not shown). The radiolabelled antibody PAb416 sticks to the vast excess of immunoglobulin protein used in the immunoprecipitation, thus generating a band of relative molecular mass 55,000 in all samples. This band is further identifiable as IgG using either radiolabelled protein A or rabbit anti-mouse IgG.

Methods. Tissues were flash frozen in liquid nitrogen, stored at -70 °C, and then homogenized in 50 mM Tris pH 7.2, 5 mM MgCl₂, 1 mM CaCl₂, 10 mM dithiothreitol, 1% NP40, 100 μ g ml⁻¹ DNase I, 50 μ g ml⁻¹ RNase A, 75 μ g ml⁻¹ phenylmethylsulphonyl fluoride. Protein concentrations were determined using the Bio-Rad protein assay reagent. 500 μ g of tissue protein or 25 μ g of cos cell protein was incubated overnight at 4 °C with 100 μ l of tissue culture supernatant containing the monoclonal antibody PAb419, and then precipitated with 100 μ l of 3% protein A-Sepharose (Pharmacia) for several hours. The immunoprecipitate was fractionated on a 10% SDS-polyacrylamide gel, transferred to nitrocellulose by electroblotting overnight, incubated first in a blocking solution (Fig. 4) for 12 h and then in a binding solution (Fig. 4) containing 10⁵ c.p.m. ml⁻¹ of PAb416, which had been radiolabelled with ¹²⁵I using the chloramine-T method. The blots were washed in Net-Gel and exposed with intensifying screens onto Kodak XAR film.

insulin/large-T antigen hybrid genes is tissue specific to the pancreas. Taken together with the immunohistochemical analysis, the transferred insulin sequences are specifying correct expression in the β -cells of the endocrine pancreas.

Penetrance and the onset of oncogenesis

The relationship between large-T antigen expression and the development of β -cell tumours bears on possible mechanisms for oncogenesis. The two previous examples of oncogene expression in transgenic mice give contrasting results²⁷. An apparently dormant SV40 early region is activated to express at high levels, and frequently amplified, in the development of choroid plexus tumours¹⁹. In contrast, a murine mammary tumour virus *myc* fusion gene is expressed in all mammary glands and in several other tissues before oncogenesis, which occurred in only 1 of the 10 mammary glands, implying that a secondary event is required to elicit tumour development²⁰. In the insulin/large-T antigen transgenic mice examined to date, a similar pattern emerges. No more than 4 or 5 of the ~100 islets in a pancreas develop into solid β -cell tumours, which are highly vascularized and readily distinguishable from normal pancreatic tissue (Fig. 3B). Yet most islets show hyperplasia and all islets examined have increased β -cell density. Furthermore, the β -cells seem to be expressing the hybrid genes well before the development of hyperplasia and tumours, as is demonstrated in an examination of the pancreas from an 8-week-old mouse (M633) in the *RIP1-Tag2* lineage. This mouse exhibited normal pathology and histology, in particular showing no hyperplasia or visible tumours in the pancreas. Thin sections were prepared from the pancreas and stained with antibodies

for insulin and large-T antigen; a representative analysis of a pair of islets is shown in Fig. 6. Large-T antigen is expressed in both of these islets, as well as in all other islets examined in this pancreas. Thus, large-T antigen is expressed in the β -cells of a young mouse before emergence of obvious hyperplasia and solid tumours.

The interpretation of these observations is that the hybrid insulin/large-T antigen genes are expressed in all β -cells, thereby inducing their proliferation, which leads to initial hyperplasia followed by oncogenic transformation. The development of solid tumours occurs in only a few of the islets, which indicates that additional events are probably necessary to mediate the rapid proliferation and extensive vascularization characteristic of the β -cell tumours. It will be of interest to distinguish between two possible classes of transformation: (1) an islet-specific change which affects most of the β -cells in an individual islet; and (2) a rare cell alteration that then produces a clonal tumour arising from that one altered cell.

Cell-cell interaction

Examination of the distribution of cell types in islets of these transgenic mice shows that islets of every size class (small, normal, enlarged, tumour) are densely packed with β -cells. Furthermore, the α - and δ -cells are disordered and rare. Comparison of the glucagon immunostains of normal islets (Fig. 1) and islets of the young transgenic mouse 633 (Fig. 6) shows that the characteristic peripheral distribution of α -cells has been disrupted and that their total number is reduced. Similarly, the normal polar organization of δ -cells is disordered and the δ -cell number severely retarded. This applies to every islet examined in this pancreas, in analyses that included sections taken to include different anatomical regions. These results suggest that during biogenesis of the islets, the α - and δ -cells are either sterically excluded or specifically suppressed by the altered β -cells.

Conclusions

The results presented here demonstrate that DNA within 520 bp upstream of the rat insulin II gene is sufficient to mediate qualitatively correct tissue- and cell-type-specific expression of hybrid insulin/SV40 large-T antigen genes. Protein blotting analysis of four mice harbouring independent insertions of such genes demonstrates expression only in the pancreas among the major tissues examined. The *in situ* analyses of thin sections from the pancreas demonstrates that this expression is cell-type specific, as large-T antigen is only detected in the insulin-producing β -cells of the islets of Langerhans.

The specification of tissue- and cell-type expression by the insulin gene 5' flanking region occurs in both possible orientations when it is located upstream of the large-T antigen structural gene. Both tissue specificity and bidirectionality are qualities previously associated with transcriptional enhancer elements using transfection of cultured cells^{28,29}. It is not clear whether this flanking region carries an additional real or cryptic promoter element oriented away from the insulin gene, or whether non-specific initiation of transcription is occurring in the *RIR-Tag* hybrid gene. RNA analyses will be necessary to address these possibilities.

There are three phases to the phenotype elicited by these hybrid oncogenes: the development of disordered islets, the subsequent enlargement of those islets into obvious hyperplasia and the formation of solid tumours. In young mice, the islets are of normal size, but densely packed with β -cells. The organization of both α - and δ -cells is disrupted and numbers of each severely reduced, δ -cells being particularly rare. These results suggest that islet cells interact to maintain a proper balance, and the transgenic β -cells effectively reduce α - and δ -cell number and organization by retarding their normal cell development. This possibility motivates a similar examination of the distribution of PP cells in the islets and islet tumours of these mice, because PP cells, like δ -cells, are often found at significant levels in naturally occurring insulinomas.

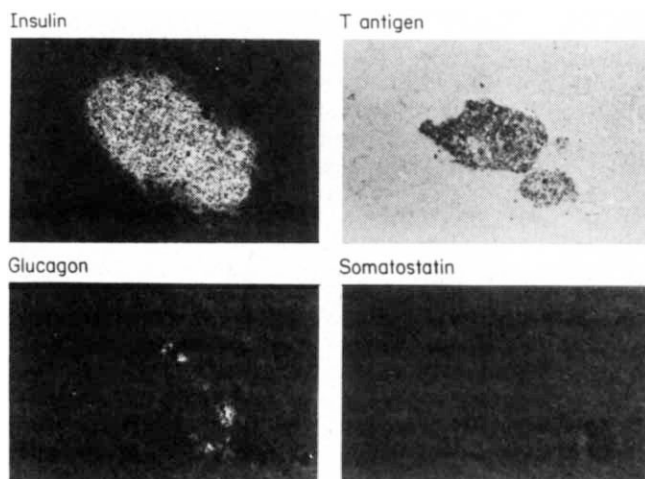


Fig. 6 Immunohistology of islets in the pancreas of a young transgenic mouse. Mouse 633, from the *RIP1-Tag2* lineage, was killed at 8 weeks of age and its pancreas flash frozen in liquid nitrogen, sectioned in a cryostat and analysed for the expression of islet cell hormones and large-T antigen as described in Fig. 4. legend. The two islets shown are representative of all islets in these and other thin sections from this pancreas. The sections used for insulin and large-T antigen are near but not adjacent and the islet sizes have changed somewhat across that distance ($\times 50$).

The prospect of disruption in islet-cell interactions is intriguing in view of observations that each of the three hormones influences secretion of the other two by their respective cell types. Somatostatin inhibits the release of both insulin and glucagon, insulin inhibits the release of glucagon and somatostatin, whereas glucagon stimulates secretion of insulin and somatostatin (see, for example, ref. 30). The organization of the endocrine pancreas into separate islands has led to suggestions that the islet cells arise from a common stem cell^{31,32}. The character of the alterations in islet composition produced by the insulin/large-T fusion gene should allow more detailed examination of islet biogenesis. If islet cell types are interacting through growth factors, the results presented here suggest that they are probably suppressors of cell growth rather than inducers. Insulin itself is one candidate for such a suppressive growth factor.

The dense packing and disorganization of islets is followed by their expansion (hyperplasia), which seems to be a direct consequence of the expression of large-T antigen. Preliminary analysis indicates that serum insulin levels are elevated (two- to ten-fold), but not in proportion to the increased number of insulin-producing cells (>100 -fold), which implies that either normal regulation of insulin secretion remains in effect, or that the β -cell structure, organization and/or vascularization is inappropriate for effective secretion. Isolation of β -cell lines from these transgenic mice should be facilitated by the ability of large-T antigen to establish (or immortalize) primary cells to growth in culture, with large-T antigen synthesis here maintained by the co-expression of the insulin and insulin/large-T antigen genes in β -cells. The availability of such β -cell lines will allow more detailed examination of the characteristics of altered β -cells. We may now examine how islet disorganization is correlated with changes in glucagon, somatostatin and insulin serum levels over a detailed timescale. It seems likely that the mice described in this study suffer from serious alterations in regulation of carbohydrate metabolism and that the explanation of their premature death does not simply result from excessive insulin in their serum, but also involves other factors.

Solid β -cell tumours eventually develop from a few of the hyperplastic islets. Tumour development occurs between 10 and 20 weeks of age, is heritable and has occurred in all progeny of the two lines (*RIP1-Tag2* and *RIR-Tag*) in which mice have lived to this age, and is occurring now in the *RIP1-Tag3* line. The solid tumours are frequent in the sense they arise in every

mouse harbouring these genes, although they occur only in a fraction of the islets. The characteristics of tumour formation suggest that synthesis of large-T antigen is necessary but insufficient to produce that condition in every islet. Large-T antigen is of a class of oncogenes that can both immortalize primary cells and release (or transform) cultured cells from their normal growth controls, such as contact inhibition and dependence on serum growth factors²⁴. Yet, *in vivo* SV40 virus will produce tumours in mice only after a long latent period; oncogenesis seems to be related to the efficacy of the immune response and the presence of large-T antigen on the cell surface as well as in the nucleus³³⁻³⁵. There is no evidence of an immune response to hyperplastic islets, nor of large-T antigen on the surface of β -cells; it remains to be established whether the frequency of solid tumour formation is associated with an escape from immune surveillance. A similar hybrid gene, composed of the 5' flanking region of the rat elastase gene linked to the coding information for SV40 large-T antigen, produces tumours of the exocrine pancreas in transgenic mice (R. Brinster and R. Palmiter, personal communication), which further indicates that T-antigen is oncogenic in mice. Given that large-T antigen synthesis is the primary event, there are other potential secondary alterations that may be necessary to induce the development of solid β -cell tumours. These include activation or inactivation of growth factors or receptors, or the cooperation of another oncogene^{36,37}. It is well established that nascent tumours cannot exceed certain size constraints without becoming vascularized; thus the induction of angiogenesis³⁸ is probably one (if not the only) secondary event necessary for solid tumour development. Biochemical and further histochemical comparisons of the islets

before and after the predictable development of tumours may provide insight into these possibilities.

The ability to examine the pre- and postnatal development of the endocrine pancreas as well as heritable oncogenesis, using as markers the hormone gene products of these cell types and the SV40 tumour antigen, will provide a means to examine the consequences of cell-specific oncogene expression and the ontogeny of the effects reported here, as well as to address further aspects of the control of insulin-gene expression. It will be of interest to compare different recombinant insulin/oncogenes (such as *myc* and *ras*) for their effects on islet-cell biogenesis and cell-cell interaction, and on secondary events necessary for the development of β -cell tumours.

I thank David Lane, Sue Hockfield, Beth Friedman and Ed Harlow for advice on immunohistology and protein blotting; Jim Watson, Winship Herr, Terri Grodzicker, Pam Mellon and Joe Sambrook for comments on the manuscript; Liz Lacy and Frank Costantini for instruction in microinjection; Stanley Minkowitz and Bartol Matanic for several pathology examinations; David Lane and Robert Santerre for antisera; Rich Cate and Wally Gilbert for the rat insulin gene; Debbie Lukralle for animal care; Mark Lacy for technical assistance; Mike Ockler and Dave Green for artwork; Marilyn Goodwin for preparation of the manuscript; and Jim Watson and Joe Sambrook for support and encouragement. D.H. was initially a Junior Fellow of the Society of Fellows, Harvard University and is currently a Special Fellow of the Leukemia Society of America. This work was funded by the Robertson Research Fund of the Cold Spring Harbor Laboratory and by a grant from Monsanto Company.

Received 18 January; accepted 20 March 1985.

- Steiner, D. F. & Freinkel, N. (eds) *Endocrine Pancreas. Handbook of Physiology* Vol. 1 (Williams & Wilkins, Baltimore, 1972).
- Cooperstein, S. J. & Watkins, D. (eds) *The Islets of Langerhans. Biochemistry, Physiology, and Pathology* (Academic, New York, 1981).
- Ullrich, A. *et al. Science* **196**, 1313-1319 (1977).
- Villa-Komaroff, L. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3727-3731 (1978).
- Lomedico, P. *et al. Cell* **18**, 545-558 (1979).
- Cordell, B. *et al. Cell* **18**, 533-543 (1979).
- Ullrich, A. *et al. Science* **204**, 612-615 (1980).
- Bell, G. I. *et al. Nature* **284**, 26-32 (1980).
- Wu, C. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1577-1580 (1981).
- Kakita, K., Goddings, S. & Permutt, M. *Proc. nat. Acad. Sci. U.S.A.* **79**, 2803-2807 (1982).
- Walker, M., Edlund, T., Boulet, A. M. & Rutter, W. J. *Nature* **306**, 557 (1983).
- Cate, R., Chick, W. & Gilbert, W. *J. biol. Chem.* **258**, 6645-6652 (1983).
- Episkopou, V., Murphy, A. & Efstratiadis, A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4657-4661 (1984).
- Palmiter, R., Chen, H. & Brinster, R. *Cell* **29**, 701-710 (1982).
- Palmiter, R. *et al. Nature* **300**, 615 (1982).
- Lacy, E., Roberts, S., Evans, E. P., Burtenshaw, M. D. & Costantini, F. D. *Cell* **34**, 343-358 (1983).
- McKnight, G., Hammer, R. E., Kuenzel, E. A. & Brinster, R. L. *Cell* **34**, 335-341 (1983).
- Brinster, R. *et al. Nature* **306**, 332 (1983).
- Brinster, R. *et al. Cell* **37**, 367-379 (1984).
- Stewart, T., Pattengale, P. & Leder, P. *Cell* **38**, 627-637 (1984).

- Grosschedl, R., Weaver, D., Baltimore, D. & Costantini, F. *Cell* **38**, 647-658 (1984).
- Swift, G., Hammer, R. E., MacDonald, R. J. & Brinster, R. L. *Cell* **38**, 639-646 (1984).
- Chada, K. *et al. Nature* **314**, 377-380 (1985).
- Toozee, J., ed. *Molecular Biology of Tumor Viruses Pt 2* (Cold Spring Harbor Laboratory, New York, 1981).
- Chick, W. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 628-632 (1977).
- Gazdar, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 3519-3523 (1980).
- Hanahan, D. *Nature* **312**, 503-504 (1984).
- Banerji, J., Rusconi, S. & Schaffner, W. *Cell* **27**, 299-308 (1981).
- Moreau, P. *et al. Nucleic Acids Res.* **9**, 6047-6068 (1981).
- Raskin, P. in *The Islets of Langerhans* (eds Cooperstein, J. J. & Watkins, D.) 467-481 (Academic, New York, 1982).
- Pictet, R. & Rutter, W. J. in *Endocrine Pancreas. Handbook of Physiology* Vol. 1 (eds Steiner, D. F. & Fienkel, N.) 25-66 (Williams & Wilkins, Baltimore, 1972).
- Pearse, A. C. E. *Nature* **295**, 96-97 (1982).
- Hargis, B. J. & Malkiel, S. J. *nat. Cancer Inst.* **63**, 965-967 (1979).
- Abrahamczuk, J., Pan, S., Maul, G. & Knowles, B. B. *J. Virol.* **49**, 540-548 (1984).
- Lewis, A. M. & Look, J. L. *Science* **227**, 15-20 (1985).
- Land, H., Parada, L. F. & Weinberg, R. A. *Nature* **304**, 596-609 (1983).
- Ruley, H. E. *Nature* **304**, 602-606 (1983).
- Folkman, J., Merler, E., Abernathy, C. & Williams, G. J. *exp. Med.* **133**, 275-288 (1971).
- Towbin, H., Staehelin, T. & Gordon, J. *Proc. natl. Acad. Sci. U.S.A.* **76**, 4350-4354 (1979).
- Burnette, W. N. *Analyt. Biochem.* **112**, 195-203 (1981).
- Harlow, E., Crawford, L. V., Pim, D. C. & Williamson, N. M. *J. Virol.* **39**, 861-869 (1981).
- Gluzman, Y. *Cell* **23**, 175-182 (1981).

LETTERS TO NATURE

Colour, albedo and nucleus size of Halley's comet

Dale P. Cruikshank*, W. K. Hartmann†
& David J. Tholen*

* Institute for Astronomy, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

† Planetary Science Institute, 2030 E. Speedway, Tucson, Arizona 85719, USA

Halley's comet (1982 i) is now bright enough for broadband photometry by large infrared telescopes¹⁻³. We report here on photometry of the comet in the *B* (0.44 μ m), *J* (0.55 μ m), *V* (1.25 μ m) and *K* (2.2 μ m) broadband filters during a time when the coma was very weak and presumed to contribute negligibly to the broadband photometry. The *V-J* and *J-K* colours suggest that the colour of the nucleus of Halley's comet is similar to that of the

D-type asteroids, which in turn suggests that the surface of the nucleus is of albedo < 0.1.

The first detection of Halley's comet in the near-infrared was made on 20 December 1984, when the comet was 5.39 AU from the Sun¹. Our new observations at *J* and *K* were made on February 18.3 and 20.3, 1985 with the 3-metre NASA Infrared Telescope Facility (IRTF) at Mauna Kea, Hawaii, and measure reflected sunlight, not thermal emission from the comet itself. Simultaneously, we obtained *B* and *V* filter data with the 2.2-metre University of Hawaii telescope, also on Mauna Kea. All data were corrected for extinction and system response by comparison with standard stars. Table 1 gives the results of the observations, together with the comet's distances from the Earth and the Sun, and the solar phase angles. The *B* and *V* photometry was done with a 9.5-arc s diameter aperture, with sky measurements of the same field taken after the comet had moved completely away. With the IRTF we used an 8-arc s aperture and a 20-arc s amplitude oscillation of the telescope's secondary mirror to correct for the sky radiation. Both telescopes were set to follow the apparent motion of the comet across the sky. We

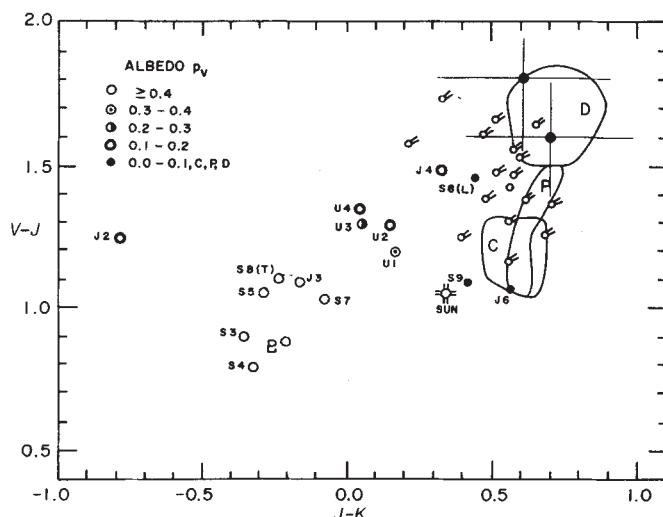


Fig. 1 *VJK* colour-colour plot of numerous planetary satellites, asteroids and comets. J, S and U refer to satellites of Jupiter, Saturn and Uranus, respectively. The fields for C, P and D-type asteroids are shown schematically, based on our data for 6 Cs, 5 Ps and 6 Ds. The two data points with error bars shown in the D field are the results for Halley's comet reported in the present work. This figure is reprinted, with modifications, from ref. 7.

could not see the comet visually through either telescope; it was located by means of a two-body ephemeris (computed by D.J.T.). The 2.2-metre telescope was then peaked on the signal through the *B* filter where the signal was strongest. The IRTF was then directed to the comet by precision offsets from field stars (observed by D.J.T.).

The question of how much light is contributed in the various filter bands by the weakly developed coma around Halley is an important factor in the interpretation of the photometry reported here. On 24 January 1985, a very weak coma was observed in 3-min unfiltered exposures with a charge-coupled device (CCD) on the 2.2-metre telescope (C. R. Chapman, D.P.C. and J. M. Pasachoff). The coma was ~ 5 arcs in diameter and was also very weak in long exposures (24 min) with a filter centred on the H_2O^+ emission band at $7,025 \text{ \AA}$. We do not have a corresponding CCD image of the comet at the time of our February photometry.

The *V* filter used in this work extends from about $5,200$ to $6,000 \text{ \AA}$ (20% transmission points) and misses most of the emission in the C_2 Swan band at $5,150 \text{ \AA}$. It includes all of the weaker C_2 band at $5,550 \text{ \AA}$. The *B* filter includes the C_3 band at $4,030 \text{ \AA}$ and C_2 at $4,700 \text{ \AA}$. All of these bands are prominent in bright, active comets. The *J* filter includes the CN (0—0) band in very bright comets, such as West (1976 vi)⁴. A preliminary evaluation of spectrophotometric observations obtained in February by A. L. Cochran (personal communication) showed no significant emission of CN or C_2 , thus suggesting that our *V* filter measurements are not seriously contaminated by coma emission bands. Similarly, the *J* filter is not expected to be contaminated because of the weakness of the CN (0—0) band. We note that any emission in the *V* filter would serve to lessen the *V-J* colour. Thus, the *V-J* values

plotted in Fig. 1 are lower limits, and the comet was at least as 'red' as we show.

Table 1 gives the observed magnitudes and $J(1, 0)$, the *J* mag reduced to unit distance from Earth and Sun and corrected to zero phase, assuming a phase coefficient of $0.039 \text{ mag deg}^{-1}$, a value commonly used for asteroids in the *V* filter. We also include a nearly contemporaneous measurement of *J* and *K* from Brooke and Knacke². In the present discussion, the colours *V-J* and *J-K* are of most interest. For February 18.3, we find $V-J = 1.6 \pm 0.2$ and $J-K = 0.7 \pm 0.3$. The mean of the three independent determinations of the *B-V* colour is 0.6 ± 0.1 .

The interpretation of these colours in the context of the small bodies of the outer Solar System (OSS) requires a brief discussion of previous work on this subject. In a series of recent papers⁵⁻⁷, we reported photometry of OSS bodies with emphasis on the colours of ices and soils on their surfaces. On bright objects, various ices (water, methane and sulphur dioxide) and mineral components of soils have been identified by medium-resolution spectrophotometry. Broadband filter observations can be used in conjunction with the spectrophotometry of the brighter objects to establish correlations between *VJK* colours and spectral classes (related to surface compositions). These correlations are assumed to apply to the fainter bodies for which only *VJK* photometry can be obtained at an adequate signal-to-noise ratio with presently available instrumentation. Halley's comet falls into this latter class because of its faintness at the solar distance when we observed it. At much smaller solar distances, the coma could be expected to develop to the point where the nucleus itself is no longer visible and the measurements would refer instead to the material in the coma.

Among small bodies in the OSS we have found a range of *VJK* colours from relatively blue to red, with clumping at the extremes, as shown in Fig. 1. The satellites of Jupiter, Saturn and Uranus (keyed with J, S and U, respectively) are known from spectrophotometry to be dominated by water ice, with increasing ice abundance towards the lower left part of the figure. Satellite J4 (Callisto) has the weakest ice bands in its spectrum. The D, P and C-type asteroids define three fields in the *V-J* and *J-K* colour diagram, as indicated by the enclosed areas in the figure.

The trend in the colours shown in Fig. 1 can be interpreted if OSS bodies are composed of mixtures of ice and dark (carbonaceous?) condensates. The surfaces, then, are mixtures of bright ices and dark soil, and the ratio of the two components determines colour and albedo. The ice/soil ratios are determined in turn by geological processes on the bodies. The *VJK* colour diagram is remarkably effective for discriminating between the C and D subtypes of dark materials. The redder D-type (abundant among Trojan asteroids) is isolated in the upper right part of Fig. 1, while the more neutral C- and P-types fall in the right centre.

The C-type asteroids in the outer belt are believed to resemble carbonaceous meteorites containing up to some 20% water. They are black objects with albedos of ~ 0.04 and flat spectra. The D-type objects have similar low albedo but a more reddish-black colour. Gradie and Veverka⁸ concluded that the D-type objects are coloured by organic material representing a lower temperature condensate than the carbonaceous material colour-

Table 1 Photometry of Halley's comet

	Date	<i>r</i> (AU)	Δ (AU)	β	<i>B</i>	<i>V</i>	<i>J</i>	<i>K</i>	<i>J</i> (1, 0)
Birkett <i>et al.</i> ¹	20.5 Dec. 1984	5.39	4.42	2.1°	—	—	18.6 ± 0.1	—	11.6 ± 0.1
Hanner and Tokunaga ³	18.3 Jan. 1985	5.12	4.30	6.6°	—	—	18.4 ± 0.2	—	11.4 ± 0.2
Brooke and Knacke ²	17.3 Feb. 1985	4.84	4.45	11.2°	$20.22 \pm 0.05^*$	$19.50 \pm 0.06^*$	18.5 ± 0.2	17.8 ± 0.3	11.9 ± 0.2
Present work	18.3 Feb. 1985	4.83	4.46	11.3°	20.18 ± 0.05	$19.56 \pm 0.07^\dagger$	17.8 ± 0.3	17.2 ± 0.3	10.6 ± 0.3
Present work	19.3 Feb. 1985	4.82	4.47	11.4°	20.08 ± 0.08	19.66 ± 0.09	—	—	—
Present work	20.3 Feb. 1985	4.81	4.48	11.5°	$20.32 \pm 0.06^\ddagger$	19.66 ± 0.09	18.1 ± 0.2	17.4 ± 0.2	11.0 ± 0.2

* *B* and *V* measurements at the 2.2-m telescope (by D.J.T.).

† Derived from *B* measured on this night and *B-V* determined on February 17, 19 and 20.

‡ Average of two measurements: at 0648 UT, $B = 20.28 \pm 0.05$; at 0842 UT, $B = 20.35 \pm 0.06$.

ing the C-types. This would be consistent with formation of the D-type objects at a great heliocentric distance. Gradie and Veverka predicted that the D-type material ought to dominate the silicate component in the outermost Solar System and might comprise the dusty constituent in comets⁸. Our observations of several comets, some of which were active and some of which were not, confirm that comets always lie near the D, P and C-class asteroid colours, as shown in Fig. 1. No comet observed by us has a colour consistent with clean ice, and the majority of them have colours just offset from the D colours. The ices of comets are therefore coloured by dark dust similar in colour to D-class asteroid material, in excellent agreement with the Gradie-Veverka prediction⁸ that comets contain D-type dust.

We note that the mean $B-V$ colour of Halley's comet is also consistent with D-type asteroids⁹. Because the comet is near to the brightening curve of an asteroid, and because of the minimal coma and emission in our wavelength regions, we may be seeing the nucleus. If so, the nucleus may have the appearance of a D-type asteroid, with a radius of about 10 km (calculated from equation 2 in ref. 13, using albedo 0.04). Alternatively, the nuclear condensation we observed may have been a swarm of particles. If those particles were macroscopic in size, colouring effects due to scattering by particles smaller than the wavelength of our observations should be minimal, and the albedos of the individual fragments would be of the order of 0.04.

Table 1, particularly the column for $J(1,0)$, shows that the reduced brightness of the comet was changing. While this could be attributed to activity in a developing coma, rather significant variations in brightness were detected at solar distance 8 AU where no comatic activity is expected¹⁰. The variations in bright-

ness have been widely considered as evidence for rotation of an irregularly shaped nucleus^{11,12}, although no unambiguous period has thus far been established.

The authors were visiting astronomers at the Infrared Telescope Facility, which is operated by the University of Hawaii under contract from the National Aeronautics and Space Administration. We thank Joseph Piscitelli for assistance at the telescope. This work was supported in part by NASA grant NGL 12-001-057. The Planetary Science Institute is a division of Science Applications International Corporation. This is PSI Contribution No. 207.

Note added in proof: Hughes¹⁴ has estimated the disk and other properties of Halley's comet from a consideration of its performance during the 1910 apparition. He finds that $D^2 p_V = 5.90 \pm 0.33 \text{ km}^2$, where D is the diameter and p_V is the geometric albedo at the V filter wavelength. The diameter derived by Hughes is approximately half that found by us at a similar albedo.

Received 25 March; accepted 10 April 1985.

1. Birkett, C., Green, S., Longmore, A. & Zarnecki, J. *IAU Circ. No. 4025* (1984).
2. Brooke, T. Y. & Knacke, R. *IAU Circ. No. 4041* (1985).
3. Hanner, M. & Tokunaga, A. T. *IAU Circ. No. 4034* (1985).
4. Johnson, R., Fink, U. & Larson, H. *Astrophys. J.* **270**, 769-777 (1983).
5. Hartmann, W. K., Cruikshank, D. P. & Degewij, J. *Icarus* **47**, 33 (1981).
6. Hartmann, W. K., Cruikshank, D. P. & Degewij, J. *Icarus* **52**, 377-408 (1982).
7. Hartmann, W. K., Cruikshank, D. P. & Tholen, D. J. in *Ices in the Solar System* (eds Benest, D., Klinger, J., Dolphus, A. & Smoluchowski, R.) (Reidel, Dordrecht, in the press).
8. Gradie, J. & Veverka, J. *Nature* **283**, 840-842 (1980).
9. Degewij, J. & van Houten, C. J. in *Asteroids* (ed. Gehrels, T.) 417-435 (University of Arizona Press, Tucson, 1979).
10. Jewitt, D. & Danielson, G. E. *Icarus* **60**, 435-444 (1984).
11. LeFevre, O. et al. *Astr. Astrophys.* **138**, L1 (1984).
12. West, R. M. & Pederson, H. *Astr. Astrophys.* **138**, L9-L10 (1984).
13. Morrison, D. & Cruikshank, D. P. *Space Sci. Rev.* **15**, 641-739 (1974).
14. Hughes, D. W. *Mon. Not. R. astr. Soc.* **213**, 103-109 (1985).

On the size of the galactic centre compact radio source: diameter $< 20 \text{ AU}$

K. Y. Lo*, D. C. Backer†, R. D. Ekers‡, K. I. Kellermann‡, M. Reid§ & J. M. Moran‡

* Owens Valley Radio Observatory, California Institute of Technology, Pasadena, California 91125, USA

† Radio Astronomy Laboratory, University of California, Berkeley, California 94720, USA

‡ National Radio Astronomy Observatory, Edgemont Rd, Charlottesville, Virginia 22903, USA

§ Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, Massachusetts 02138, USA

The existence of a compact nonthermal radio source at the galactic centre was first suggested by Lynden-Bell and Rees¹ as the possible signature of a black hole. Such a source was detected unambiguously by Balick and Brown². Very long baseline interferometry (VLBI) observations of the structure of this source, hereafter referred to as Sgr A*, have been made for several years^{3,4}. In all previous experiments, the visibility amplitudes, measured over a restricted (U, V) range, could be fitted only to the simplest model of brightness distribution, that is, to a circular gaussian. We have re-observed Sgr A* at $\lambda 3.6 \text{ cm}$ and $\lambda 1.35 \text{ cm}$ with many baselines, using the more sensitive receivers and VLBI recording terminals now available. These observations set an upper limit of 20 AU ($3 \times 10^{14} \text{ cm}$) to the diameter of Sgr A* at $\lambda 1.35 \text{ cm}$ and reveal for the first time an elongated structure at $\lambda 3.6 \text{ cm}$, with the position angle of the long axis almost parallel to the rotation axis of the Galaxy. Sgr A* is unique in our Galaxy, but resembles most closely the compact radio sources at the centre of external galaxies. Observations of the central 4 arc s (0.2 pc) at radio and other wavelengths are best explained by a single massive collapsed object at the galactic centre.

The $\lambda 3.6\text{-cm}$ observations were made on 7 May 1983, using six telescopes: Haystack 36-m; National Radio Astronomy Observatory (NRAO) 43-m; Fort Davis 26-m; Goldstone 64-m (DSS14); Owens Valley Radio Observatory (OVRO) 40-m; and Hat Creek (HCRK) 26-m. The MkIII VLBI recording terminals⁵ were used with an effective recording bandwidth of 28 MHz. Observations alternated between Sgr A* and the neighbouring

calibrator source NRAO530. The absolute flux-density scale was fixed by observations of NRAO530 relative to 3C286 (assumed flux density of 5.4 Jy) with DSS14.

Observations at $\lambda 1.35 \text{ cm}$ were made on 23 June 1983, using four telescopes: Haystack 36-m, NRAO 43-m, Very Large Array (VLA) phased D-array and OVRO 40-m. The MkIII recording terminal was used with a 56-MHz bandwidth. We calibrated with the H_2O maser source in Sgr B2, $\sim 0.7^\circ$ away. At the end of each Sgr A* observation, Sgr B2 was observed for 1 min. We have adjusted the scale so that the total flux density of Sgr A* was $\sim 1.5 \text{ Jy}$, a value consistent with the measurement at the VLA during the observations. The size of the gaussian model discussed below is completely independent of the flux-density scale.

At $\lambda 3.6 \text{ cm}$, the calibrator source NRAO530 was detected on all baselines, whereas the compact radio source in Sgr A* was detected only on the short baselines formed between the DSS14, OVRO and HCRK telescopes (for visibility coverage, see ref. 6). The visibility amplitudes of Sgr A* (Fig. 1) were fitted to two simple models, namely, circular and elliptical gaussian brightness distributions. The baseline most sensitive to the difference of the two models is DSS14-HCRK because of the relatively large rotation of the baseline. The elliptical model is preferred.

The elliptical gaussian model based on all three baselines has the following parameters: the major-axis (half-power) diameter is $15.5 \pm 0.1 \text{ milliarc s}$ (marcs), and an axial ratio of 0.55 ± 0.25 , with the long axis oriented at a position angle (PA) of $98 \pm 15^\circ$. From 08:30 to 10:00 UT, the closure phase of Sgr A* on the DSS14-OVRO-HCRK triangle is $< 5^\circ$. This suggests that the elongated source is very symmetrical, but the closure phase is measured over a very limited time.

The present $\lambda 3.6\text{-cm}$ results, together with our earlier work³ spanning 8 yr, set an upper limit of $0.6 \text{ marcs yr}^{-1}$ to any change in the observed size (corresponding to a radial expansion velocity of $\sim 16 \text{ km s}^{-1}$ at 10 kpc). The combined data on the transcontinental baselines to DSS14 set an upper limit of 0.011 Jy (5σ) at $\lambda 3.6 \text{ cm}$ for any component of Sgr A* with diameter $\leq 1 \text{ marcs}$, within a $2 \text{ arc s} \times 2 \text{ arc s}$ box centred on $\text{RA}(1950) = 17 \text{ h } 42 \text{ min } 29.3 \text{ s}$, $\text{dec.}(1950) = -28^\circ 59' 18.2''$.

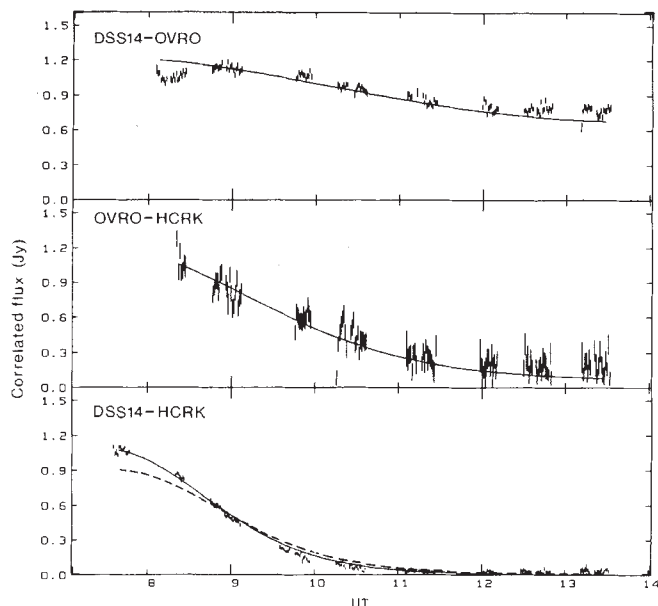


Fig. 1 Fringe amplitudes of the galactic centre compact radio source Sgr A* at $\lambda 3.6$ cm (7 May 1983). The values beyond 11:00 UT on the HCRK baselines correspond to 2.5σ upper limits to the correlated flux density. The error bars represent $\pm 1\sigma$ in the fringe fitting. The solid curves represent the best-fitting elliptical gaussian model visibility (major-axis half-power diameter = 15.5 ± 0.1 marc s, PA = $98 \pm 15^\circ$, axial ratio = 0.55 ± 0.25). The dashed curve is the circular gaussian model visibility. On the other two baselines, the circular and elliptical gaussian model visibilities are essentially the same.

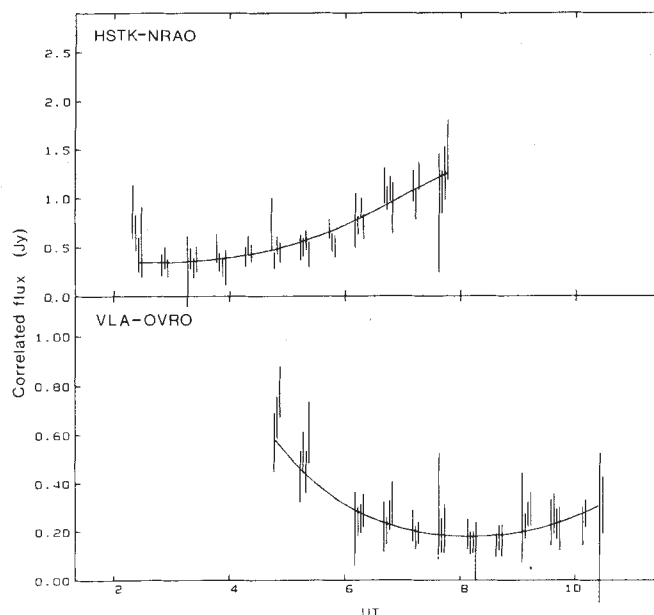


Fig. 2 Fringe amplitudes of Sgr A* at $\lambda 1.35$ cm (23 June 1983). The solid curves are the elliptical gaussian model visibilities (major-axis diameter = 2.2 ± 0.2 marc s, PA = $87 \pm 30^\circ$, axial ratio = 0.55 ± 0.5).

The $\lambda 1.35$ -cm data (Fig. 2) are well fitted by a circular gaussian model with a diameter of 2.1 ± 0.3 marc s. An elliptical gaussian fit, starting with the above model, results in a major-axis diameter of 2.2 ± 0.2 marc s, a position angle of $87 \pm 30^\circ$ and an axial ratio of 0.55 ± 0.5 . The $\lambda 1.35$ -cm data are thus consistent with the elongated morphology observed at $\lambda 3.6$ cm, scaled by λ^2 , but do not require it.

The wavelength dependence of the angular size of Sgr A* can be derived from these data, the previously published 6-cm data³ and 13-cm observations⁷. These are the only four wavelengths for which the angular size has been determined from visibility data with a range of variation exceeding a factor of two. All observations are mainly sensitive to the east-west brightness distribution. The sizes at $\lambda = 13, 6, 3.6$ and 1.35 cm are 203, 49.0, 15.5 and 2.1 marc s, respectively. Although these observations were obtained at different times, there is no evidence for time variations of the angular size. The exponent of the power-law dependence of the size on wavelength is 2.0 ± 0.1 .

The λ^2 dependence of the source diameter could be attributed to self-absorption by internal thermal electrons^{3,8}, or ascribed to angular broadening of the intrinsic source size resulting from scattering by interstellar electrons⁹. The observed elongation of Sgr A* does not preclude the scattering interpretation, as a symmetrical elongated structure can be caused by scattering within a confined layer of anisotropic turbulence^{10,11}. Asymmetric scattering from anisotropic turbulence in the interplanetary medium near the Sun has recently been observed (B. Rickett, personal communication). A brightness-distribution map of Sgr A* could resolve this ambiguity, because a complicated source structure would preclude interstellar scattering. Under either interpretation, the $\lambda 1.35$ -cm source size of 2.1 marc s (~ 20 AU) is an upper limit to the intrinsic source size of Sgr A*. Unless there is an abrupt change in the λ^2 dependence of the observed size for $\lambda < 1.35$ cm, the true size is ≤ 1.5 marc s.

If the observed size is dominated by scattering, the upper limit to the source-expansion rate observed at $\lambda 3.6$ cm may not apply to the intrinsic expansion. However, even in the scattering case, the small $\lambda 1.35$ -cm size places an upper limit on the intrinsic source expansion of 0.26 ($2.1/8$) marc s yr⁻¹, or

~ 6 km s⁻¹, for the intrinsic source size not to exceed the scattering angle. Therefore, measurements at the two wavelengths imply an upper limit to the intrinsic source expansion velocity of < 16 km s⁻¹ averaged over 8 yr.

The observed parameters of Sgr A* are summarized in Table 1; its properties are similar to those of compact radio sources in galactic nuclei and quasars¹², but at a much lower luminosity. The physical conditions of the radio source implied by the observed properties have been discussed previously^{8,13}. The extremely small size indicates that the scale of the underlying energy source must approach stellar dimensions. However, models of the radio source invoking stellar objects, such as a young supernova remnant, a pulsar⁹, pulsar-driven wind¹³ or a binary stellar radio source⁸, are inconsistent with the observations³. Furthermore, there is no other radio source with comparable strength to Sgr A* found in the central five square degrees of the galaxy (D.C.B. and R. Sramek, unpublished results). Sgr A* is a unique radio source in the Galaxy. On the other hand, it resembles closely a weaker version of the nuclear radio source in external spiral galaxies, such as M81 (NGC3031; ref. 14) and M104 (NGC4594; ref. 15).

Problems of confinement of the radiating particles¹³ and the stability of the source size would require the emission to originate in a steady wind or jet. The observed elongation of the source that is almost parallel to the rotation axis of the Galaxy (PA $\sim 122^\circ$) is possible supporting evidence. Further observational constraints on the nature of the radio source must await the construction of a brightness-distribution map.

Table 1 Observed properties of the galactic centre compact radio source

Source size	< 20 AU ($\sim 3 \times 10^{14}$ cm)
Wavelength dependence of size	$\lambda^{2.0}$ ($\lambda \geq 1.35$ cm)
Position angle of elongation	$98 \pm 15^\circ$ ($\lambda = 3.6$ cm)
Axial ratio	0.55 ± 0.25 ($\lambda = 3.6$ cm)
Upper limit to source expansion	< 16 km s ⁻¹
Flux density variability ($\Delta S/2S$)*	≤ 0.2 (1975-83)
Spectral index (ref. 29)	~ 0.25
Turnover frequency (ref. 29)	≥ 89 GHz
Radio luminosity	$\sim 2 \times 10^{34}$ erg s ⁻¹
Brightness temperature	$> 7 \times 10^8$ K ($\lambda 1.35$ cm)

* ΔS , peak-to-peak variation; S , mean flux density; ref. 28 and J. van Gorkom, K.Y.L. and M. Claussen, unpublished results.

The central few square degrees of the Galaxy are full of interesting and energetic phenomena on many size scales¹⁶⁻¹⁹, in the middle of which lies the unique Sgr A*, coincident to within 0.5 arcs of IRS16 centre. IRS16 may be the primary source of the $\sim 10^7 L_{\odot}$ of infrared and ionizing radiation from the centre²⁰⁻²². A $3 \times 10^6 M_{\odot}$ point mass within 0.5 pc of the centre is suggested by the latest Ne II observations (E. Serabyn and J. H. Lacy, in preparation). Furthermore, a large velocity ($\pm 750 \text{ km s}^{-1}$) outflow^{23,24} and, possibly, a compact source of positronium²⁵ are also identified with the central 0.2 pc ($\sim 5 \times 10^{17} \text{ cm}$, or $\leq 4 \text{ arc s}$).

Although the luminosity and ionization over the 10^{20}-cm scale of the galactic centre may be explained by a recent burst of star formation²⁶, the activities within the central $5 \times 10^{17} \text{ cm}$ require an alternative source of energy, such as a massive collapsed object in a quiescent state of low-level accretion²⁷. The unique compact radio source and the other phenomena at the centre would be natural consequences of the presence of such an object.

We thank M. Cohen for encouragement and for reviewing the paper critically and Alan Rogers and Paul Ste Marie for help in the data reduction. The VLBI programs at OVRO, Hat Creek, Haystack and Fort Davis are supported by the NSF. NRAO is operated by the Associated Universities Inc. under contract with the NSF.

Received 7 February; accepted 13 March 1985.

- Lynden-Bell, D. & Rees, M. J. *Mon. Not. R. astr. Soc.* **152**, 461-475 (1971).
- Balick, B. & Brown, R. L. *Astrophys. J.* **194**, 265-270 (1974).
- Lo, K. Y., Cohen, M. H., Readhead, A. C. S. & Backer, D. C. *Astrophys. J.* **249**, 504-512 (1981).
- Kellermann, K. I., Sahaffier, D. B., Clark, B. G. & Geldzahler, B. G. *Astrophys. J. Lett.* **214**, L61-L62 (1977).
- Rogers, A. E. E. *et al. Science* **219**, 51-54 (1983).
- Lo, K. Y. *IAU Symp.* **110**, 265-273 (1984).
- Marcaide, J. M. *et al. in Conference on Active Galactic Nuclei* (ed. Dyson, J. E.) (Manchester University Press, in the press).
- Brown, R. L., Lo, K. Y. & Johnston, K. J. *Astr. J.* **83**, 1594-1597 (1975).
- Davies, R. D., Walsh, D. & Booth, R. *Mon. Not. R. astr. Soc.* **177**, 319-333 (1976).
- Higdon, J. C. *Astrophys. J.* **285**, 109-123 (1984).
- Cordes, J. M., Ananthakrishnan, S. & Dennison, B. *Nature* **309**, 689-690 (1984).
- Kellermann, K. I. & Pauliny-Toth, I. I. K. *A. Rev. Astr. Astrophys.* **19**, 373-410 (1984).
- Reynolds, S. P. & McKee, C. F. *Astrophys. J.* **239**, 893-897 (1980).
- Bartel, *et al. Astrophys. J.* **262**, 556-563 (1982).
- Hummel, E., van der Hulst, J. M. & Dickey, J. *Astr. Astrophys.* **134**, 207-221 (1985).
- Sofue, Y. & Handa, T. *Nature* **310**, 568-569 (1984).
- Yusef-Zadeh, F., Morris, M. & Chance, D. *Nature* **310**, 557-561 (1984).
- Lo, K. Y. & Claussen, M. J. *Nature* **306**, 647-651 (1983).
- Ekers, R. D., van Gorkom, J., Schwarz, U. J. & Goss, M. W. *Astr. Astrophys.* **122**, 143-150 (1983).
- Becklin, E., Gatley, I. & Werner, M. *Astrophys. J.* **258**, 135-142 (1982).
- Lacy, J. H., Townes, C. H. & Hollenbach, D. J. *Astrophys. J.* **262**, 120-134 (1982).
- Henry, J. P., Depoy, D. L. & Becklin, E. E. *Astrophys. J. Lett.* **285**, L27-L30 (1984).
- Hall, D. N. B., Kleinman, S. G. & Scoville, N. Z. *Astrophys. J. Lett.* **262**, L53-L58 (1982).
- Geballe, T. R. *et al. Astrophys. J.* **284**, 118-125 (1984).
- Ramaty, R. & Lingenfelter, R. E. *Highlights of Astronomy* **6**, 525-529 (1983).
- Rieke, G. & Lebofsky, M. J. *AIP Conf. Proc.* **83**, 194-203 (1982).
- Rees, M. *AIP Conf. Proc.* **83**, 166-176 (1982).
- Brown, R. L. & Lo, K. Y., *Astrophys. J.* **253**, 108-114 (1982).
- Backer, D. C. *IAU Symp.* **97**, 389-390 (1982).

Middle atmospheric NO and NO₂ observed by the Spacelab grille spectrometer

J. Laurent*, M.-P. Lemaitre*, J. Besson*, A. Girard*, C. Lippens†, C. Muller†, J. Vercheval† & M. Ackerman†

*Office National d'Études et de Recherches Aéronautiques, F-92322 Châtillon, France

†Belgium Institute for Space Aeronomy, B-1180 Brussels, Belgium

Almost 20 years after Nicolet¹ suggested the presence of nitric oxide in the upper atmosphere as a source of D-region ionization by solar Lyman α -radiation, this trace species was first observed² by the resonance fluorescence method; almost 10 years later, its vertical distribution was observed in the stratosphere by means of balloon-borne infrared absorption spectrometry³. Since then much

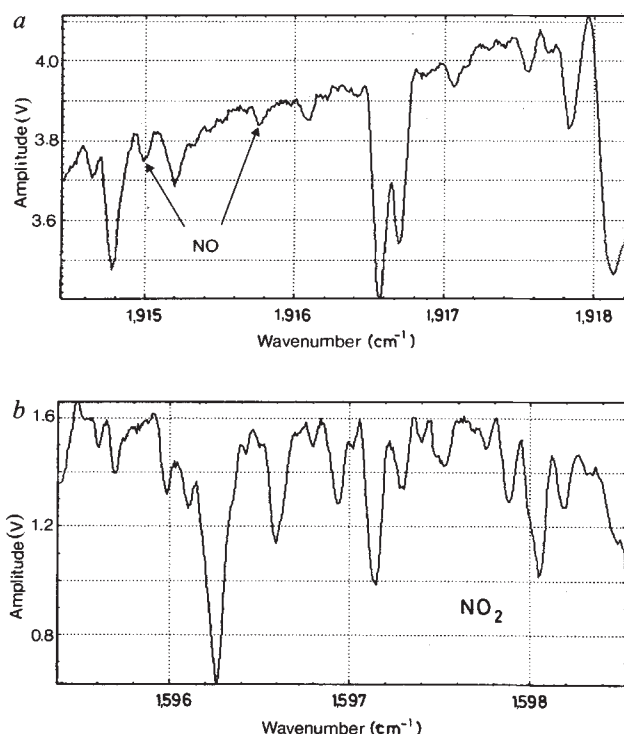


Fig. 1 *a*, Solar spectrum recorded at sunlight-tangent altitude equal to 86 km. Two NO absorption lines are indicated; all the other spectra lines are attributable to solar absorption features. *b*, NO₂ absorption spectrum recorded at sunlight-tangent altitude equal to 26 km. In this spectral region, the solar spectrum is a continuum. Most of the spectral features are attributable to nitrogen dioxide. In *a* and *b*, amplitudes are in arbitrary units and the origin of the scale is the level of total absorption for the solar radiation.

emphasis has been placed on odd nitrogen compounds in the stratosphere, partly because of their role in controlling the ozone abundance. Many measurements have been made of NO, NO₂, HNO₃ and NO₃ in the stratosphere with inferences on the N₂O₅ abundance. Meanwhile, several determinations of NO in the mesosphere and in the lower thermosphere⁴⁻¹⁰ have used resonance fluorescence and mass spectrometry¹¹. We report here the first observation of NO from the low thermosphere down to the low stratosphere by instrumentation similar to that used previously on board balloon gondolas, but on this occasion the observation platform was Spacelab I, which gave access to higher altitudes. As before¹², NO and NO₂ were observed simultaneously.

The infrared grille spectrometer on board the space shuttle during the first Spacelab mission has been described elsewhere¹³. During this mission, one Earth-limb solar occultation run was devoted to the observation of NO and NO₂ in the respective wavenumber ranges, 1,914.5-1,918 cm⁻¹ and 1,595-1,598.5 cm⁻¹. The corresponding wavelengths were selected in the sixth and fifth orders of the grating by two interference filters, one for each of the two liquid nitrogen temperature-cooled detectors.

This observation took place over the southern Pacific Ocean on 1 December 1983, at sunrise after the very short nights of the high southern latitudes in their late spring. For the 20- and 100-km tangent-altitude locations, the solar depression angles did not exceed 0.3° and 1.4° respectively, during the course of the day. The geographical coordinates of the tangent points of the solar rays in the atmosphere were respectively 68° S, 129° W and 67° S, 118° W. Thus, the observations took place in almost full daylight; almost all the odd nitrogen was in the form of NO in the mesosphere and NO, NO₂ and HNO₃ in the stratosphere, according to our present knowledge of atmospheric photochemistry. NO and NO₂ spectra (examples are shown in Fig. 1*a, b*) were recorded every 2 s and the measured equivalent

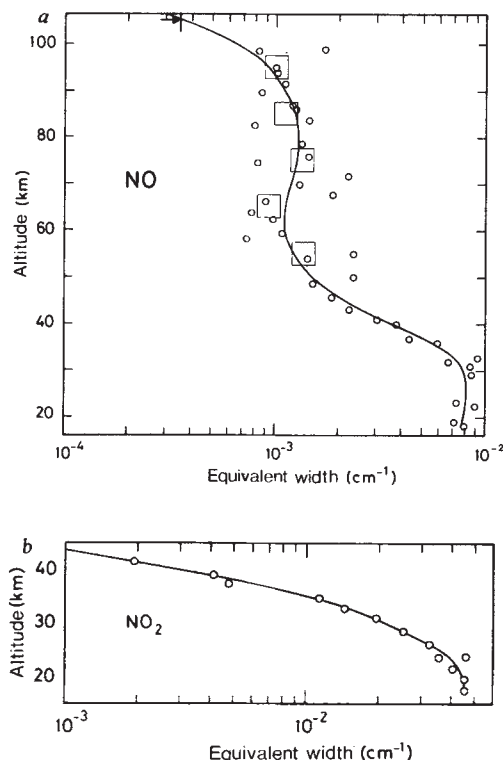


Fig. 2 Equivalent widths versus grazing-sunlight altitude for *a*, the NO absorption doublet at $1,914.99\text{ cm}^{-1}$ (the squares represent five kilometres averages) and *b*, the NO_2 multiplet centre at $1,597.15\text{ cm}^{-1}$.

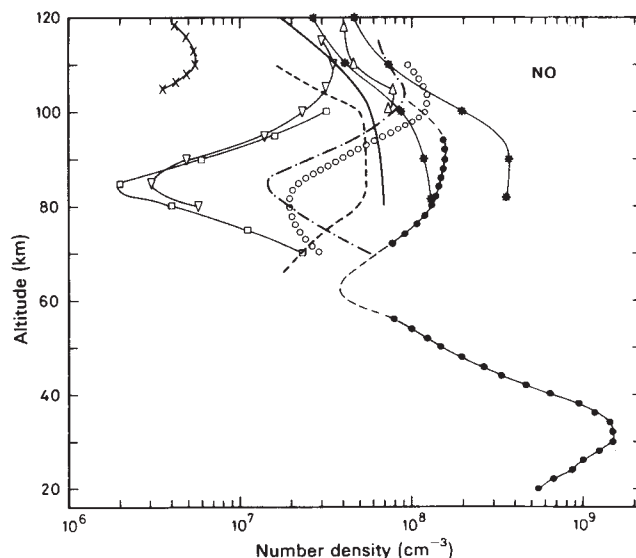


Fig. 3 Number densities of NO and NO_2 versus altitude (1 December 1983). The near straight lines indicate constant volume mixing ratios from 10^{-10} to 10^{-5} . Error bars are related to the scattering of the measured equivalent widths.

altitude a dashed line represents inverted NO number densities, so low that the slant absorption must be attributed to NO at higher altitudes. Hence, at this altitude the NO abundances can only be given tentatively and probably represent an upper limit.

The stratospheric portion of the observation is characterized by a high signal-to-noise ratio for the absorption lines. Hence, the uncertainty on the volume concentrations is small and the values, which will be discussed elsewhere, are in the range of previous observations. Our results are compared in Fig. 4 with other observations in the mesosphere and in the low thermosphere. Although our results are in rather good agreement with other observations at the highest altitudes, the differences become very large at 85 km, particularly with the vertical profiles exhibiting a minimum volume concentration at that altitude. The shape of the vertical profile compares as well as possible with the results of Barth² and Tisone⁵. The only other absorption measurements to compare with our data are those of Massie¹⁰, based on OSO-8 observations of the (1.0) δ band at 182.94 nm, which are in good agreement.

Slight modifications of our new data may be necessary if further refinement were to alter significantly the orbit parameters of the shuttle during the first Spacelab mission. More observations will be possible when the instrument is flown in a mission dedicated to atmospheric studies. When this happens we should gain understanding of the the distribution of NO in the mesosphere, which shows a disagreement between the NO abundances derived from different observation methods. This has important consequences for modelling neutral and charged-particle abundances in the upper-middle atmosphere and ionospheric D and E regions.

Received 22 January; accepted 19 March 1985.

1. Nicolet, M. *Mem. Inst. R. Met. Belg.* **19**, 1-143 (1945).
2. Barth, C. A. *J. Geophys. Res.* **69**, 3301-3303 (1964).
3. Ackerman, M. *et al. Nature* **245**, 205-206 (1973).
4. Meira, L. G. *J. geophys. Res.* **76**, 202-212 (1971).
5. Tisone, G. C. *J. geophys. Res.* **78**, 746-750 (1973).
6. Witt, G., Dye, J. E. & Wilhelm, N. *J. atmos. terr. Phys.* **38**, 233-238 (1976).
7. Tohmatsu, T. & Iwagami, N. *Space Res.* **15**, 241 (1975).
8. Baker, K. D. *et al. J. Geophys. Res.* **82**, 3301-3303 (1964).
9. Thomas, R. J. *J. geophys. Res.* **83**, 513-516 (1978).
10. Massie, S. T. *J. geophys. Res.* **85**, 2155-2164 (1980).
11. Trinks, H., von Zahn, U., Barth, C. A. & Kelly, K. K. *J. geophys. Res.* **83**, 203-206 (1978).
12. Ackerman, M. *et al. Planet. Space Sci.* **23**, 651-660 (1975).
13. Lemaître, M.-P. *et al. Science* **225**, 171-172 (1984).
14. Rothman, L. S. *et al. Appl. Opt.* **22**, 2247 (1983).

widths versus tangent heights in the atmosphere are shown in Fig. 2a, b.

The data, represented by the smooth curves in Fig. 2a, b, were invented using the molecular-line parameters of Rothman *et al.*¹⁴ in the 'onion peeling' inversion method, which leads to the vertical distributions of NO and NO_2 shown in Fig. 3. Above 98 km altitude, the scanned wavelength intervals broadened to observe broader band solar spectra. The $1,914.99\text{-cm}^{-1}$ NO line region was observed again at 105 km tangent altitude where the NO signal was too weak to be detected. From 55-75 km of

Measurement of surface currents in Liverpool Bay by high-frequency radar

D. Prandle & D. K. Ryder

Institute of Oceanographic Sciences, Bidston Observatory, Birkenhead, Merseyside L43 7RA, UK

The backscattering of high-frequency radar echoes from the sea surface can yield useful information about ocean waves and surface currents¹⁻³. The Rutherford Appellton Laboratory Ocean Surface Current Radar (OSCR) system is based on the CODAR system⁴ but differs significantly in that it uses a beam-forming receive antenna to determine the direction of the backscattered signals instead of using the CODAR direction finding system. OSCAR derives a radial component of surface velocity by measuring the shift in recorded spectra associated with advancing and receding waves. In this survey of surface currents in Liverpool Bay, using only OSCAR, measurements were made first for 7 days from the Lancashire coast followed, 1 month later, by similar measurements from the Wirral coast. These were combined to compute current ellipses for the predominant M_2 tidal constituent. Spatial distributions of the ellipse parameters agree well with values obtained from a numerical model and from moored current meters. This new observational technique should provide data on the fine structure of near-shore surface currents for use in studies of sediment transport and pollutant dispersal.

The handicap of having only one OSCAR unit was overcome by deploying it in an area where tidal currents predominate. Tidal measurements from two separate deployments could be combined to produce a stereoscopic image. Liverpool Bay was selected as a suitable site because, in addition to the convenient shoreline geometry (Fig. 1), the region is characterized by strong tidal currents flowing over complex bottom topography, thus producing horizontal structure on a scale commensurate with the resolution of OSCAR. Despite the influence of major estuaries, the tidal forcing is sufficient to overcome vertical stratification. Moreover, harmonic analysis of tidal elevations in this region indicates the predominance of three major constituents, namely

M_2 , S_2 and N_2 . This allows a reasonable resolution of the primary M_2 constituent from a data set as short as 25 h.

The computing system installed in OSCAR at the time of this experiment limited data acquisition to measurements along just four beams per hour. Hence, the net duration of data recorded for each beam was restricted to between 1 and 4 days. Measurements were made over 14 beams each 6° wide; resolution along each beam consisted of 20 bins each 1.2 km long. Previous preliminary trial results⁵ have indicated that, using an integration period of about 10 min, current measurements accurate to 1 cm s^{-1} can be obtained.

A simplified tidal analysis was performed by correlating (1) the time series $Z(t)$ for tidal elevations at Liverpool (predicted from harmonic constants) with (2) the time series $U(t)$ for the radial current component in a specific bin. A phase lag δt was introduced between these time series and the error $\epsilon(\delta t)$ determined for the relationship

$$\epsilon(\delta t) = \sum_{i=1}^N [(U(t) - \bar{U}) - \alpha(Z(t + \delta t) - \bar{Z})] \quad (1)$$

where N is the total number of relevant observations, $\bar{U}$ and $\bar{Z}$ are means of the respective time series and the coefficient α was calculated by the method of least squares to minimize ϵ .

In this manner, the value for $\delta t_{\min}$ (within the range -12 h to $+12 \text{ h}$) corresponding to the smallest overall value of ϵ was determined. The value $\delta t_{\min}$ was then assumed to represent the phase lag of current relative to elevation for the predominant M_2 constituent. Similarly the amplitude of the M_2 constituent of tidal current was assumed to be given by the product of α (corresponding to $\delta t_{\min}$) and the M_2 amplitude for tidal elevation. The accuracy of this analysis was aided by the quiescent meteorological conditions with winds < 10 knots for 80% of the time.

At each intersection point of the two beam sets shown in Fig. 1, the elliptical pattern of surface tidal currents for the M_2 constituent was calculated by linear interpolation of the respective current components in the nearest two bins. The component parts of Fig. 2 indicate the spatial distribution of the four parameters concerned: (1) amplitude of the major axis of the ellipse; (2) direction of the major axis; (3) phase or timing of the occurrence of maximum current relative to the lunar transit; and (4) ellipticity, with positive values indicating anti-clockwise rotation. Figure 2 also indicates comparable values

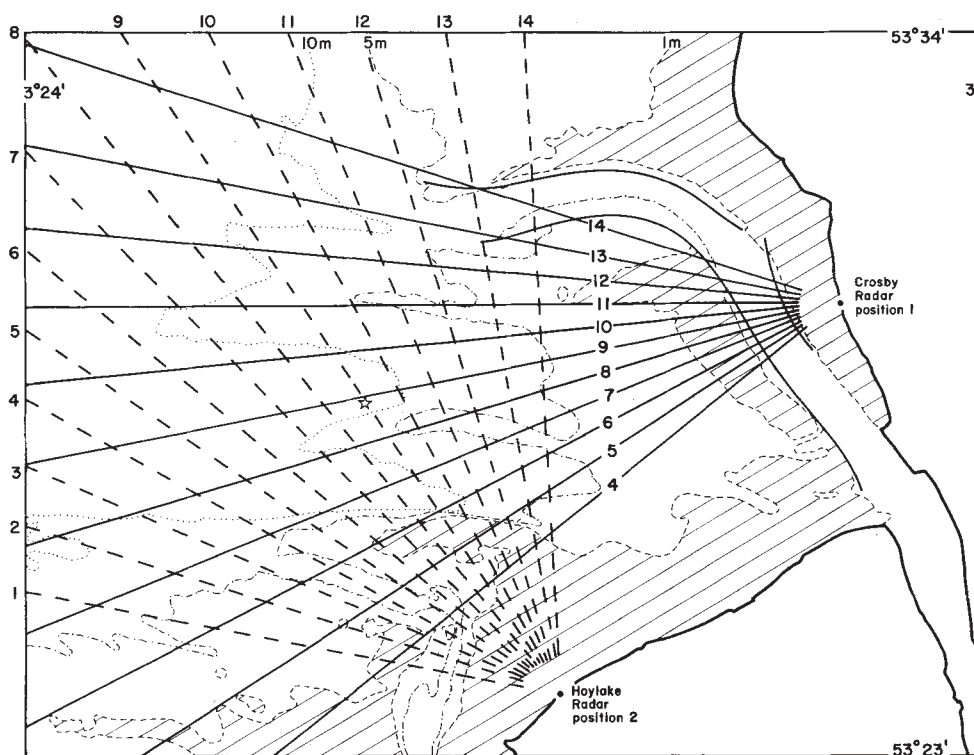


Fig. 1 Liverpool Bay showing centre lines of radar beams. ☆, Location of moored current meters.

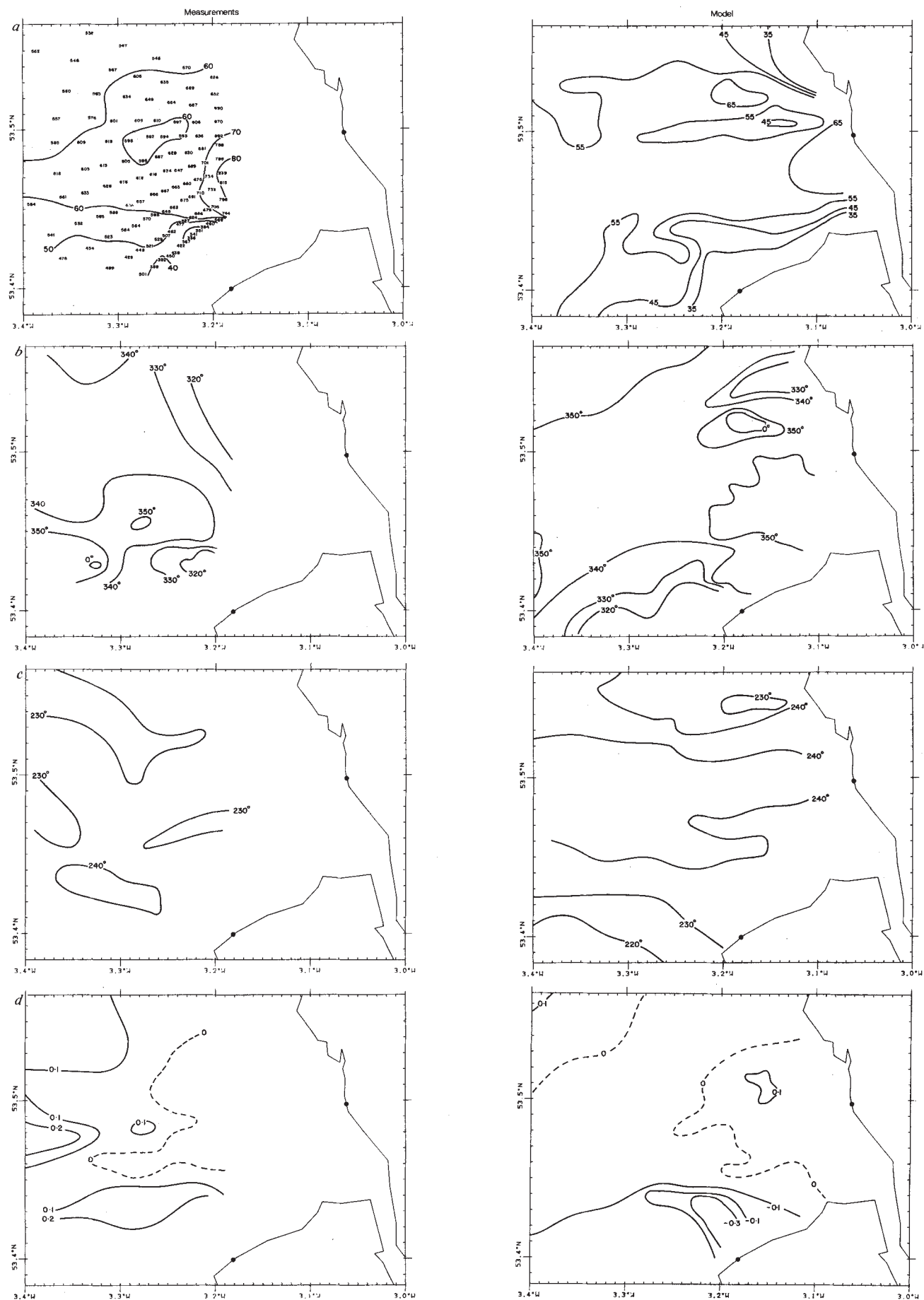


Fig. 2 M_2 current ellipse parameters from surface measurements by high-frequency radar (left), and vertically integrated numerical model (right). *a*, Amplitude of the major axis in cm s^{-1} ; *b*, direction of major axis, measured anticlockwise from the east; *c*, phase, time of maximum current after lunar transit at Greenwich; *d*, eccentricity, minor axis/major axis, +ve anticlockwise rotation.

Table 1 M_2 tidal ellipse properties at 53° 28.3'N, 3° 15.4'W

	Amplitude (major axis) (cm s^{-1})	Direction* (deg)	Phase (deg)	Eccentricity (+ve anti- clockwise)
Numerical model	58	350	243	-0.02
Current meters (1)	54	351	237	-0.02
(2)	50	356	231	-0.04
High-frequency radar (surface)	61	342	232	+0.03
High-frequency radar corrected to mid- depth value	55	344	229	+0.08

* Measured anticlockwise from due east.

obtained from a numerical model based on a 1-km rectangular grid. In the model, the tides are specified at both the northern and southern limits of the Irish Sea and comparison between tide gauge recordings and model computations for M_2 elevations in Liverpool Bay indicates agreement to within 2% in amplitude and 7° in phase. Figure 2a is reproduced to a larger scale to illustrate the actual values of current amplitude obtained from the radar observations. The spatial coherence indicated by these values suggests that the precision of the radar measurements is of the order of 1 cm s^{-1} , thus substantiating theoretical estimates. Moreover, the agreement between spatial patterns from the radar measurements and the model establishes the accuracy of this instrument.

The surface amplitudes indicated by the radar are larger than the depth-mean values calculated in the model. This tendency is predicted from theoretical analysis of current structure^{6,7}. Thus, for a position in the centre of the survey region, Table 1 indicates M_2 ellipse properties from (1) the numerical model, (2) two current meters, (3) the high-frequency radar surface measurements and (4) the high-frequency radar measurements corrected to reflect conditions at mid-depth. (The current meter deployments formed part of the OSCAR evaluation experiment.) These latter corrections were carried out by resolving the ellipse into clockwise and anticlockwise components and assuming the current structure shown in Fig. 5 of ref. 6. Relative to surface values, the current ellipse at mid-depth exhibits a smaller amplitude, phase advance, increase in anticlockwise eccentricity and the direction may veer in either sense⁷.

Table 1 also indicates that for the two current meters, located just 500 m apart, amplitude may vary by 10% and both phase and direction by 5°. Thus, allowing for the inherent spatial variability in current ellipse properties, both vertically and horizontally, the small discrepancies shown in Table 1 between the ellipse properties measured by the high-frequency radar and those obtained elsewhere appear entirely explicable. Likewise, the agreement between the high-frequency radar measurements and the numerical model shown in Fig. 2b, c and d for direction, phase and eccentricity respectively is within the limits of expected variability.

The residual circulation in Liverpool Bay is influenced by tides, winds and density gradients and hence may be highly variable in time. To construct a residual flow pattern from the present radar data involving velocity components measured on separate days up to 5 weeks apart must be of questionable validity. However, an examination of the residual current data was made both as an illustrative example of the potential use of dual OSCAR systems and to indicate the order of magnitude of surface residuals even if the spatial distribution proved incoherent. Residual currents R for each bin were obtained from the expression

$$R = \bar{U} - \alpha \bar{Z} \quad (2)$$

where $\bar{U}$ and $\bar{Z}$ represent the means calculated in equation (1) and the value of α corresponds to $\delta t_{\min}$. Then for each intersection point of the two beam sets, a residual current vector was calculated using linear interpolation between the respective

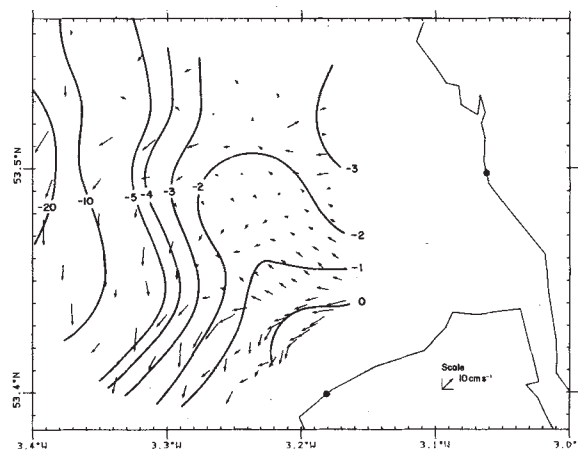


Fig. 3 Residual circulation measured by high-frequency radar. Streamlines are in units of $10^3 \text{ m}^3 \text{ s}^{-1}$.

residual current components in the nearest two bins. Figure 3 shows these observed residual current vectors together with superimposed streamlines. The streamlines are based on the assumption that the surface residuals persist uniformly throughout the water column. The streamlines were constructed using the intersection of beam 9 from Hoylake and beam 5 from Crosby as a zero datum. The spacing of the streamlines was then determined along each beam from Hoylake by calculating the total flow orthogonal to each beam relative to a datum along beam 5 from Crosby. These datum values for beam 5 (Crosby) were initially determined in similar fashion by calculating flow orthogonal to the beam.

The results shown in Fig. 3 suggest residual currents at the surface of up to 19 cm s^{-1} with a mean value of about 5 cm s^{-1} . The maximum currents occur in the deeper water on the western boundary of the survey region. In this region, a strong southerly flow is indicated involving a flow of $20 \times 10^3 \text{ m}^3 \text{ s}^{-1}$, this latter circulation may be compared with a freshwater discharge of the Mersey of $\sim 10^2 \text{ m}^3 \text{ s}^{-1}$ and a net flow through the Irish Sea of about $60 \times 10^3 \text{ m}^3 \text{ s}^{-1}$.

Over the shallower regions in the eastern section, the residuals are much smaller and form an anticlockwise gyre. In the south-western section, a pronounced westerly circulation is suggested. Residuals measured by the moored current meters averaged over the period 1 April to 17 May 1984 were 5.6 cm s^{-1} towards the south-southwest 4.3 cm s^{-1} eastwards. The latter may reflect the proximity of this mooring to the knoll shown in Fig. 1 while the former value reflects asymmetry in the direction of flood and ebb currents of some 10° . The magnitude of these currents is of the same order as the surface residuals shown in this region in Fig. 1, while the variation in their directions illustrates the limitations of conventional instruments for measuring residuals in regions of this kind.

Despite extensive efforts to measure the residual circulation in Liverpool Bay⁸ there is no clear pattern against which to evaluate Fig. 3.

One residual flow component, seemingly overlooked in these earlier studies, is the Stokes drift arising from variation in total water depth over the tidal cycle. For the predominant M_2 constituent, a tidal elevation $\hat{Z} \cos \omega t$ combined with a tidal current $\hat{U} \cos (\omega t - g)$ produces a residual flow $\frac{1}{2} \hat{U} \hat{Z} \cos g$. Using the distribution of M_2 currents shown in Fig. 2 with the $\hat{Z}$ value of (295 cm, 318°) we calculate a net flow over the survey region due to Stokes drift of $10 \times 10^3 \text{ m}^3 \text{ s}^{-1}$ directed southeastwards towards the Mersey (the direction corresponds to the net propagation of tidal energy). The velocities associated with this flow range from 2 cm s^{-1} in the deepest water to over 10 cm s^{-1} in the shallower southeasterly region. This Stokes transport must be counterbalanced by an Eulerian residual current component. The latter requirement may partially explain the strong westerly flow observed in the southern region of the radar data. However,

both observations and theory suggest that residual circulation in areas of the kind considered differ significantly between surface and bed and such considerations are beyond the scope of the present paper.

The residual circulation thus forms a coherent pattern with current magnitudes comparable with theoretically derived residual flow components. Most excitingly, high-frequency radar offers the possibility of measuring the residual circulations (at the surface) in such regions where conventional instruments have proven ineffective.

This initial evaluation of OSCAR has confirmed the accuracy of the instrument and provided unique data on both tidal and residual circulation in Liverpool Bay. Further, more rigorous, determination of the accuracy of OSCAR measurements is hindered by the absence of comparable independent measurements, the present results suggest a significant advance over the levels of accuracy cited in earlier high-frequency radar measurement⁹⁻¹¹.

Using the two OSCAR units now available, similar surveys in other coastal areas can, in as little as 7 days, produce current data which may well supplant conventional data sets obtained over many years of observations. For the first time, the fine detail of spatial structure of surface currents can be measured. This should improve our theoretical understanding of the circulation within bays, around headlands, sand banks and so on, and thus be of direct benefit in studies of sediment transport, dispersal of pollutants or ship navigation.

The OSCAR system was developed at the Rutherford-Appleton Laboratory under the direction of J. W. King; other members of the development group, F. D. G. Bennett, D. Eccles and K. Slater, participated directly in the Liverpool Bay experiment. Current meter observations were analysed by J. Howarth (IOS Bidston). We thank other staff from Bidston who participated in the measurement programme. D.K.R. is in receipt of a post-graduate studentship from the NERC.

Received 4 December 1984; accepted 5 March 1985.

1. Barrick, D. E. *Boundary-Layer Met.* **13**, 23-43 (1978).
2. Shearman, E. D. R. *Proc. Inst. elect. Engr.* **130**, 579-590 (1983).
3. Leise, J. A. *IEEE J. Ocean Engng* OE-9, 106-113 (1984).
4. Barrick, D. E., Evans, M. W. & Webber, B. L. *Science* **198**, 138-144 (1977).
5. King, J. W. et al. *Current Measurements Offshore* (Society for Underwater Technology, London, 1984).
6. Prandle, D. *Continental Shelf Res.* **1**, 191-207 (1982).
7. Prandle, D. *Geophys. Astrophys. Fluid Dyn.* **22**, 29-49 (1982).
8. *Out of Sight Out of Mind* Vol. 1-4 (HMSO, London, 1972-76).
9. Frisch, A. S. & Weber, B. L. *J. geophys. Res.* **85**, 485-493 (1980).
10. Holbrook, J. R. & Frisch, A. S. *J. geophys. Res.* **86**, 10908-10912 (1981).
11. Maresca, J. W. Jr, Padden, P. R., Chang, R. T. & Seibel, E. *Limnol. Oceanogr.* **25**, 929-935 (1980).

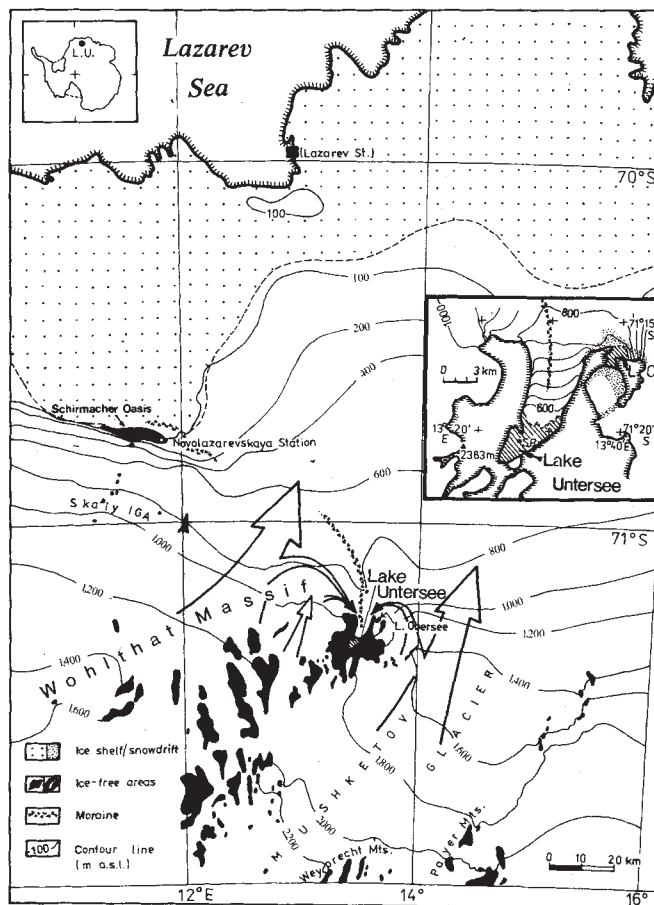


Fig. 1 Sketch map of the working area with locations mentioned in the text. The arrows show the main glacier flows feeding Lake Untersee. L.U., Lake Untersee; L.O., Lake Obersee; S.P., sampling point.

stantly losing water through sublimation on the surface of the more than 2.5-m-thick ice cover. The salt content suggests that the present water body is the remainder of an amount of melt water at least 50 times as great.

On the northeastern margin of the mountains in central Dronning Maud Land, pilots of the German Antarctic Expedition in 1939 discovered two large dark ice planes under which liquid water was presumed to be present—Lake Untersee and Lake Obersee³ (Fig. 1). Then, in 1969, a Soviet group explored Lake Untersee and proved the existence of a large water body under a 2.5-m-thick ice cover².

In March 1982 our group first implemented a hydrological sampling programme with the following results: (1) At the end of the summer, the water body sampled at depths of 5, 10, ..., 40, 44 m was homogeneous down to the bottom (44 m), with respect to temperature ($T_{\text{water}} = (+0.45 \pm 0.05)^\circ\text{C}$), isotope composition ($\delta^2\text{H}_{\text{water}} = (-255 \pm 5)\text{‰}$, $\delta^{18}\text{O}_{\text{water}} = (-35.0 \pm 0.5)\text{‰}$) and salt content (sum of ions) ($M_{\text{water}} = (330 \pm 10)\text{mg l}^{-1}$). Probably this remarkable homogeneity (within analytical error) results from convection, induced by the greenhouse effect of the ice cover, as known from lakes of the Schirmacher Oasis⁴. (2) The ice cover (interval 0-1.3 m) is characterized by $\delta^2\text{H}_{\text{ice}} = (-250 \pm 5)\text{‰}$ and $\delta^{18}\text{O}_{\text{ice}} = (-34.0 \pm 0.5)\text{‰}$. (The techniques of sample preparation and isotope analysis are described by Mühle et al.⁵. Isotope data refer to Standard Mean Ocean Water (SMOW).) (3) The tritium content of the water is below the detection limit, that is, <0.5 tritium units (TU). For comparison, large lakes of the Schirmacher Oasis, fed with melt water from old ice as well as from recent snow patches, have tritium contents of 6-8 TU (ref. 6). (4) Lake water is of sodium sulphate type with an extremely low magnesium content (Table 1, Fig. 2). (5) There are signs of primary production in the lake. As already observed² the bottom

Lake Untersee, a first isotope study of the largest freshwater lake in the interior of East Antarctica

W.-D. Hermichen, P. Kowski & U. Wand

Akademie der Wissenschaften der DDR, Zentralinstitut für Isotopen- und Strahlenforschung (ZFI), Permoserstrasse 15, DDR-7050 Leipzig, GDR

Perennially ice-covered lakes partially bounded by glacier ice are common in some Antarctic coastal regions¹ but very rare inside the continent. Here we describe hydrological studies on the largest freshwater lake of interior Antarctica, Lake Untersee ($71^\circ 20' \text{S}/13^\circ 30' \text{E}$, surface area 10 km^2 , maximum depth $>79 \text{ m}$)², showing that the lake arose from a melt-water pond during climatic optimum periods in the Holocene. At present, the studied water body is thermally, hydrogeochemically and isotopically homogeneous because of thermal convection during the austral summer. Lake Untersee is fed throughout the year by underwater melting of the adjoining glacier ice. Isotope data suggest a permanent ice cover during its existence. The drainless lake is con-

is at least partially covered with an organic layer a few millimetres thick.

No continuous record of meteorological data is available, but the recent climate of the Lake Untersee region may be compared with that of the South Victoria Land dry valleys⁷: mean annual temperature of $\sim -20^\circ\text{C}$ and high summer aridity producing an annual sublimation rate of $>200\text{ mm}$ of ice. In March 1982 no appreciable thawing was observed near the lake. The great difference in altitude between the lake surface (620 m above sea level (a.s.l.)) and that of the glaciers south of Lake Untersee (1,500–2,000 m a.s.l.) gives rise to strong catabatic winds.

Under recent climatic conditions the formation of such a large water body cannot be expected. Therefore, and because of the impossibility of increased endogenous heat flux, the origin of Lake Untersee must be sought in Holocene climatic optimum periods. The elucidation of the palaeoclimatic conditions of this area is the aim of investigations on organic deposits (fossilized stomach oils of petrels)⁸ and ice-cores⁹.

Relics of strandlines of boulders above Lake Untersee³ indicate a former higher level of the lake and probably represent a pause in icesheet retreat (compare ref. 1). The salt content of the lake water is high compared with that of Antarctic snow and ice from locations at similar distances from the coast or similar altitudes (see Table 1)^{10–19}. Marine aerosols (mainly Na^+ and Cl^-) and chemical weathering of rocks (mainly Ca^{2+} and SO_4^{2-}) are the principal sources of salts. Presumably, the local weathering solutions have a high Ca/Mg ratio because of widely distributed anorthosites ($\text{CaO/MgO} > 50$) in the Lake Untersee area.

For water-balance studies, chloride, which enters the lake exclusively as deposited marine aerosol via melting glacier ice, seems to be the most suitable indicator. A comparison of the Cl^- concentration of lake water (43.4 mg l^{-1}) and interior Antarctic precipitations ($0.05\text{--}1\text{ mg l}^{-1}$; see Table 1 and refs 10, 12, 13, 15, 19) shows the present-day water body to be the remainder of an at least 50 times larger quantity of melt water.

The isotope data indicate that Lake Untersee has been supplied mainly by underwater melting of old glacier ice in recent times. The hydrogeochemical characteristics of this inflow are determined by the precipitation chemistry of two nourishing areas (see Fig. 1): (1) the area between Richthofen, Weyprecht and Payer Mountains with a surface level of 1,400–2,000 m a.s.l. and deuterium contents of precipitations ranging between about -220 and -300‰ (ref. 20) and (2) the firn basin glacier between the Skaly IGA nunataks and the Wohlthat Massif with $\delta^2\text{H} \sim -200\text{‰}$ (ref. 9).

If more than 200 mm are removed annually from the ice cover

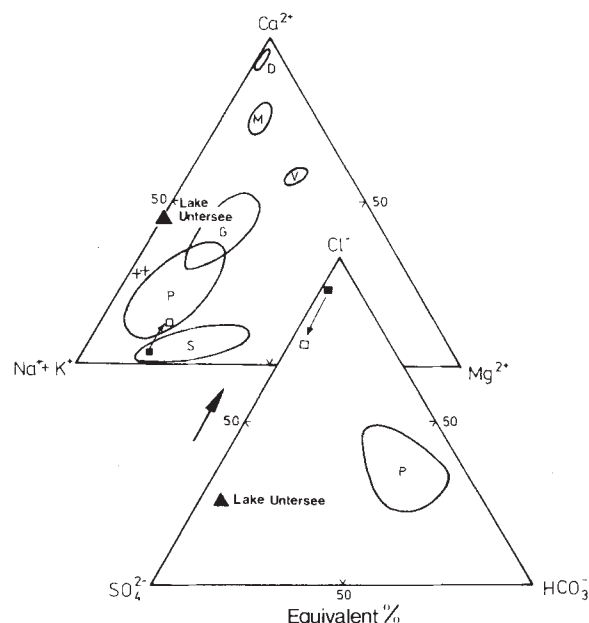


Fig. 2 Chemical composition of Lake Untersee water compared with other Antarctic waters. P, Atmospheric precipitation, Schirmacher Oasis, Dronning Maud Land (after ref. 6); D, Don Juan Pond, saline water; G, glacial melt water; M, Lake Miers, fresh water; S, Antarctic snow and ice; V, Lake Vanda, saline water (fields from ref. 28); $\blacksquare$, Antarctic sea water (after ref. 26); $\square$, average atmospheric precipitation over the oceans (after ref. 29); +, snow, Lake Untersee area (see Table 1).

of the lake by sublimation, a corresponding regeneration of the floating ice cover may occur through freezing of water onto its underside during the low insolation season. However, the observed isotope differences $\delta_{\text{ice}} - \delta_{\text{water}}$ are lower than expected for a thermodynamically-equilibrated ice/water system²¹ by $\sim 15\text{‰}$ ($\delta^2\text{H}$) and $\sim 1\text{‰}$ ($\delta^{18}\text{O}$); the same tendency was observed in some lakes of the Schirmacher Oasis⁴. In both cases the ice formation seems to be diffusion controlled: concentration of the heavy isotopes in the forming ice produces a depletion in ^2H and ^{18}O in the adjacent water. If the rate of downward growth of ice is high enough, say 0.1 cm h^{-1} , diffusion, being the only active mixing process within the stably stratified water system (diffusion coefficient $D_{\text{water}}(0^\circ\text{C}) = 1 \times 10^{-5}\text{ cm}^2\text{ s}^{-1}$)²², is unable to rehomogenize the liquid phase. Thus, the deeper freezing ice equilibrates thermodynamically not with the bulk water of the

Table 1 Chemical composition of Lake Untersee water and of Antarctic atmospheric precipitations sampled in the working area

Sample	Na^+	K^+	Ca^{2+}	Mg^{2+}	HCO_3^-	Cl^-	SO_4^{2-}	$\frac{\text{Na}^+}{\text{Cl}^-}$	$\frac{\text{SO}_4^{2-}}{\text{Cl}^-}$	$\frac{\text{Ca}^{2+}}{\text{Mg}^{2+}}$
Lake Untersee, average lake water (2.7–44 m)	60.5	4.25	44.5	0.18	16.7	43.4	158.1	1.4	3.6	247.2
Snow*, Lake Untersee area ($\sim 1,250\text{ m a.s.l.}$)	4.0	1.1	1.7	0.15		4.7	4.2	0.8	0.9	11.3
Snow* on the ice cover of Lake Untersee	2.5	0.5	0.95	0.05		2.4	3.1	1.0	1.3	19.0
Freshly fallen snow, Schirmacher Oasis ²⁴	0.2–2.5	0.07–0.6	$<0.05\text{--}0.8$	$<0.05\text{--}0.48$		1.8–3.6		0.3–0.9		0.4–10
Firn/ice, Law Dome, East Antarctica ²⁵ (distance from coast 110 km, 1,400 m a.s.l.)	0.050	0.006	0.004	0.008						0.5
Snow, James Ross Island, Antarctic Peninsula (1,660 m a.s.l.) ¹³	0.25	0.009	0.013			0.58		0.43		
Atmospheric precipitation, Mirnyy Station ¹¹ (annual mean)	0.98	0.19	0.057	0.087		1.24	0.38	0.79	0.31	0.65
Sea water ²⁶								0.56	0.14	0.04

Concentrations (mg l^{-1}) of cations were determined using an atomic absorption spectrometer; those of anions, by standard techniques of water analysis²⁷.

* Snow presumably altered by sublimation causing unusually high salt contents.

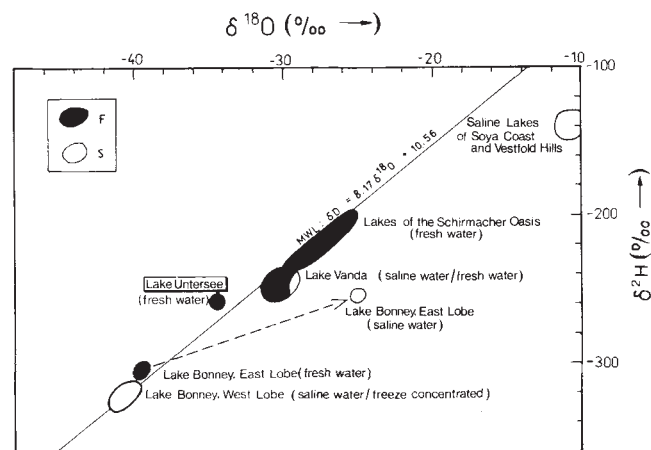


Fig. 3 Position of Lake Untersee water in the $\delta^2\text{H}/\delta^{18}\text{O}$ diagram (values versus SMOW). For comparison, data²³ of other Antarctic lakes with different evolutions are also shown. Freshwater lakes (type F) are slightly shifted isotopically to the left of the Meteoric Water Line (MWL)²¹ by freeze-concentration. Saline lakes (type S) are, as a rule, isotopically shifted well to the right by dominating evaporation-concentration.

lake but with isotopically somewhat lighter water of the boundary layer. Because the sublimation of the ice cover takes place layer by layer without any isotope separation, the heavy-isotope content of the remaining liquid water decreases as the amount of water decreases²¹ and it remains at slightly lowered values if input and sublimation are equilibrated²³. In both cases, an increase in salt concentration is noted.

Lake Untersee has been predominantly or permanently covered with ice throughout its evolution. This is demonstrated by the position of its water isotope data in the $\delta^2\text{H}/\delta^{18}\text{O}$ diagram (Fig. 3), slightly shifted to the left from the global Meteoric Water Line (MWL)²¹ by the isotope fractionation in the water/ice transition. The permanence of an ice cover is also supported by the occurrence of supra-glacial till boulders resting on the ice cover (compare ref. 1). The dominance of evaporation at an ice-free surface would have produced a shift to the right side of the MWL as observed in most of the Antarctic saline lakes²³.

A comparison of the strandline of the Lake Untersee on the basis of air photographs taken in 1939, 1961 and 1982 does not show any relevant change in the shape of the lake. Obviously, its mass balance in this period was approximately equalized.

More extensive work is in progress on the Lake Untersee environment as well as its Holocene evolution.

We thank the leaders and the helicopter crews of the 26th and 27th Soviet Antarctic Expeditions, M. Grelle for assistance in sampling, Drs M. Partisch and R. Zierath for hydro-geochemical analyses, Dr W. Kurze for tritium analyses and Mrs E. Kirschner for mass spectrometric measurements. This work was supported by the Central Institute for Physics of the Earth, Academy of Sciences of the GDR, Potsdam.

Received 6 June 1984; accepted 29 January 1985.

- Pickard, J. & Adamson, D. A. in *Antarctic Earth Science* (eds Oliver, R. L., James, P. R. & Jago, J. B.) 460-472 (Australian Academy of Science, Canberra, 1983).
- Kosenko, N. G. & Kolobov, D. D. *Byull. Sov. Antarkt. Eksped.* **79**, 65-69 (1970).
- Ritscher, A. *Wissenschaftliche und fliegerische Ergebnisse der Deutschen Antarktischen Expedition 1938/39* Vol. 1 (Leipzig, 1942).
- Richter, W. & Strauch, G. *Isotopenpraxis* **19**, 145-153 (1983).
- Mühle, K. et al. *ZfI-Mitt.* **37**, (1981).
- Richter, W., Wand, U., Strauch, G., Kowski, P. & Kurze, W. *ZfI-Mitt.* **89**, 47-73 (1984).
- Myagkov, S. M. *Byull. Sov. Antarkt. Eksped.* **100**, 99-101 (1980).
- Hiller, A. & Wand, U. *ZfI-Mitt.* **89**, 103-121 (1984).
- Hermichen, W.-D., Kowski, P. & Strauch, G. *ZfI-Mitt.* **89**, 87-102 (1984).
- Murozumi, M., Nakamura, S. & Yoshida, Y. *Mem. natn. Inst. Polar Res., spec. Iss.* **7**, 255-263 (1978).
- Melnikov, S. A., Rusanov, V. P. & Smagin, V. M. *Byull. Sov. Antarkt. Eksped.* **104**, 9-15 (1983).
- Vilenskiy, V. D. & Koroleva, N. I. *Antarktika* **13**, 147-156 (1974).
- Aristarain, A. J., Delmas, R. J. & Briat, M. *J. geophys. Res.* **C87**, 11004-11012 (1982).
- Delmas, R. J., Briat, M. & Legrand, M. *J. geophys. Res.* **84**, 4314-4318 (1982).
- Briat, M., Boutron, C. & Lorius, C. *J. Rech. Atmos.* **8**, 895-901 (1974).
- Herron, M. M. & Langway, C. C. Jr *J. Glaciol.* **24**, 345-356 (1979).
- Boutron, C. & Lorius, C. *Int. Ass. Hydrol. Sci. Pub.* **118**, 164-171 (1977).
- Warburton, J. A. & Linkletter, G. O. *J. Glaciol.* **20**, 149-162 (1978).

- Mumford, J. W. & Peel, D. A. *Brit. Antarct. Surv. Bull.* **56**, 37-47 (1982).
- Kotlyakov, V. M. & Gordienko, F. G. *Isotope and Geochemical Glaciology* (Gidrometeoizdat, Leningrad, 1982) (in Russian).
- Int. Atom. Energy Ag. Tech. Rep. Ser. No. 210*, (1981).
- Landolt-Börnstein, *Zahlenwerte und Funktionen aus Physik, Chemie, Astronomie, Geophysik, Technik* 6th ed. Vol. 5a (ed. Schäfer, K.) 405 (Springer, Berlin, 1969).
- Matsubaya, O., Sakai, H., Torii, T., Burton, H. & Kerry, K. *Geochim. cosmochim. Acta* **43**, 7-25 (1979).
- Wand, U., Hermichen, W.-D., Partisch, M. & Zierath, R. *Geodät. geophys. Veröff. R.I.* **12**, (in the press).
- Johnson, B. B. & Chamberlain, J. M. *Geochim. cosmochim. Acta* **45**, 771-776 (1981).
- Evans, A. J. *Aust. natn. Antarct. Res. Exped. Sci. Rep.* **110** (1970).
- Tümping, H. von (ed.) *Ausgewählte Methoden der Wasseruntersuchung* Vol. 1 (Gustav Fischer, Jena, 1971).
- Torii, T., Murata, S. & Yamagata, N. *J. R. Soc. N.Z.* **11**, 387-399 (1981).
- Savenko, V. S. *Geokhimiya* **12**, 1890-1893 (1976).

A new method for monitoring temporal trends in the acidity of fresh waters

Andrew L. Herczeg, Wallace S. Broecker,
Robert F. Anderson & Sherry L. Schiff

Lamont-Doherty Geological Observatory and Department of
Geological Sciences, Columbia University, Palisades,
New York 10964, USA

David W. Schindler

Freshwater Institute, 501 University Crescent, Winnipeg,
Manitoba, Canada R3T 2N6

Acid precipitation and the consequent increased acidity of fresh waters are subjects of widespread concern. Several organizations in the United States, Canada and Europe have established extensive monitoring programmes in an attempt to determine the relationship between the acidity of precipitation and the change in acidity of fresh waters. We contend here that the present strategies for monitoring temporal trends in the acidity of fresh waters will not yield results sensitive enough to detect trends in acidity even on a 10-year timescale. We propose here an alternative scheme based on calculations using precise measurements of dissolved inorganic carbon (DIC) and partial pressure of CO_2 (P_{CO_2}). This scheme eliminates biases inherent in pH electrode determinations and minimizes the perturbations in acidity caused by changes in the P_{CO_2} of lake waters associated with biological cycles.

A method that accurately and precisely yields hydrogen-ion activity is needed for the following reasons: (1) Aquatic biota are sensitive to hydrogen-ion activity. Most laboratory bioassay experiments refer to changes in, for example, quantity, distribution and viability of a given species as a function of pH. (2) Trace metal mobilization is one of the major effects of lake acidification. Hydrogen-ion activity is an essential parameter in metal speciation calculations. (3) Total alkalinity measurements cannot reliably substitute for measurements of hydrogen-ion activity. The presence of titratable organic anions leads to an overestimate of bicarbonate ion from acid titration data and hence to an underestimate of hydrogen-ion activity at any given P_{CO_2} . (4) The predicted response of a lake to acidification for a given pH of rain is complicated by many unknown factors such as dry fallout, soil and sediment neutralization and HNO_3 sequestering. Policies to set anthropogenic emissions of nitrogen and sulphur oxides must be based on sound predictive models derived from accurate empirical data.

The difficulties associated with measurement of pH using the standard glass/calomel- or combination-electrode systems have been well documented. In dilute natural waters, electrode determinations are very sensitive to sampling and handling artefacts and to liquid junction potential problems¹⁻⁶. The use of a flow-cell electrode system⁵ has to some extent remedied the last problem, although the technique has not yet received widespread application. More serious are interactions of dissolved organic carbon (DOC) and/or suspended matter with the electrodes, which cause systematic errors of up to 0.5 pH units²⁻⁴.

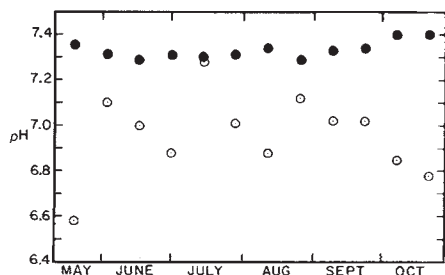


Fig. 1 Comparison of pH measured by electrodes ($\odot$) with pH calculated from DIC and P_{CO_2} data using equation (2) ($\bullet$) for ELA lake 239 during the summer of 1982. DIC can be determined by acidifying a small volume of water to $pH < 4$ and then stripping the CO_2 with helium gas. The gas phase is analysed for CO_2 on a gas chromatograph or infrared analyser to a precision of better than 1%¹¹. P_{CO_2} is determined by equilibrating an aliquot of water with a small volume of air and then an aliquot of the air is returned to the laboratory for GC or infrared analysis. Replicate equilibrations agree to within 2%. Equilibrations should be performed in the field because of the extreme sensitivity of the P_{CO_2} to biological processes. The hydrogen-ion content can thus be calculated with a precision of $\pm 4\%$ at pH 7 and $\pm 10\%$ at pH 5.6, which is substantially better than can be obtained by electrode measurements ($\pm 20\%$). The analytical precision of the pH electrode is a moot point because of the systematic errors of up to 0.5 pH units (a factor of three in H^+ activity) in the pH range 5.6–7. The measurements of DIC and P_{CO_2} are no more difficult or expensive than quality pH determination.

An intercomparison among 17 laboratories using standard electrode techniques was conducted, in which each was given an aliquot of a sample known to have a pH of 6.18 (ref. 6). They reported a mean value of 5.54 with a standard deviation of 0.74 pH units. At pH values below the accuracy and agreement were significantly better and it may be that the biases that plague electrode measurements become smaller as the hydrogen-ion activity increases.

As well as the abovementioned analytical problems, pH is sensitive to changes in P_{CO_2} (ref. 7). Photosynthesis, respiration and entrainment of hypolimnetic water, by temporarily raising or lowering P_{CO_2} , create a perturbation of the longer-term pH trend. For example, lakes that normally have a $pH \approx 7$ when at equilibrium with atmospheric P_{CO_2} may have their pH driven as high as 10 during phytoplankton blooms because of the photosynthetic consumption of CO_2 (refs. 8–10). Although equilibrating water samples with standard atmospheric P_{CO_2} (340 μatm) potentially provides a means of circumventing this difficulty, this approach is subject to the electrode errors described above.

Highly accurate and reproducible hydrogen-ion concentrations can be calculated from measurements of DIC and P_{CO_2} (ref. 4) from the equilibrium relationship

$$pH = -\log \frac{[H_2CO_3^*]K_1}{([DIC] - [H_2CO_3^*])} \quad (1)$$

where $[HCO_3^-] = [DIC] - [H_2CO_3^*]$ at $pH < 8$ and $K_1 = [HCO_3^-][H^+]/[H_2CO_3^*]$, and $[H_2CO_3^*] = [H_2CO_3] + [CO_2]$.

The analysis of P_{CO_2} and DIC allows pH to be normalized to that which would prevail for equilibrium with the ambient P_{CO_2} in the atmosphere above the lake. The normalization is carried out as follows with the assumption that alkalinity is conservative:

$$pH^0 = -\log \left[\frac{P_{CO_2}^0 K_H K_1 [H^+]_c}{[H_2CO_3^*] K_1} \right] \quad (2)$$

where the superscript refers to the value of a given parameter at mean atmospheric P_{CO_2} and 20 °C, K_H is the solubility of CO_2 and $[H^+]_c$ is calculated from DIC and P_{CO_2} measurements (equation (1)).

A comparison of pH obtained by direct measurement using a glass electrode with that calculated from DIC/ P_{CO_2} measure-

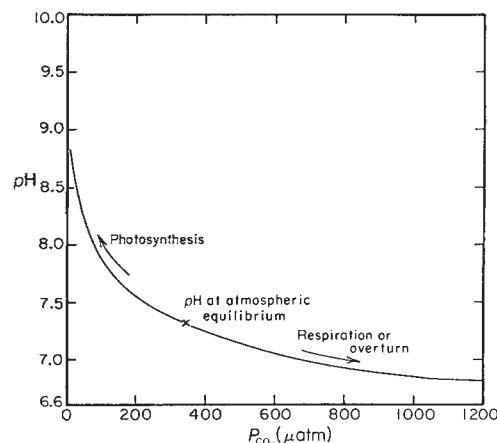


Fig. 2 Calculated dependence of pH on P_{CO_2} for lake 239 surface water. ($[HCO_3^-] = 115 \mu mol l^{-1}$ at 20 °C.) The range corresponds to that which may be caused by algal blooms or overturn events. An upper P_{CO_2} limit of $\sim 1,000 \mu atm$ can be achieved for surface water during fall overturn when hypolimnetic water ($P_{CO_2} = 3,400 \mu atm$; A.H., unpublished data) is brought to the surface at a faster rate than gas exchange can remove it to the atmosphere. Although we did not observe a P_{CO_2} depression in lake 239 during the summer, the response of eutrophic lakes to phytoplankton blooms^{7–9} suggests that the P_{CO_2} of those lakes must be well below 10 μatm .

ments is shown in Fig. 1. Lake 239 at the Experimental Lakes Area (ELA), northwestern Ontario, is a Canadian Shield lake that was routinely monitored for pH , DIC and P_{CO_2} during the summer of 1982. The dissolved CO_2 content ($H_2CO_3^*$) of surface water averaged $13.5 \mu mol l^{-1}$ (normalized to 20 °C) during this period with a standard deviation of $0.3 \mu mol l^{-1}$ and was always close to equilibrium with the P_{CO_2} of the atmosphere. The normalized pH results calculated from the DIC/ P_{CO_2} data show only a very small scatter around a mean of 7.33. In contrast, the electrode-measured pH data scatter over a range of 0.7 pH units with a mean of 6.93. The pH measurements were made using standard laboratory techniques at the ELA chemistry laboratory¹¹. The comparison illustrates both the poor precision and inaccuracy of the pH electrode.

The calculated response of lake 239 surface water to changes in P_{CO_2} is shown in Fig. 2. Extremely low or high P_{CO_2} that can occur during phytoplankton blooms or overturn events, respectively, could produce variations in pH as large as 2 pH units (100-fold H^+ activity). Therefore, even if the *in situ* pH could be electrometrically measured accurately, time trends in acidity would be hard to detect if the P_{CO_2} fluctuations were not taken into account. However, when the P_{CO_2} /DIC method is used, the calculated pH (equation (1)) can be normalized as in equation (2), thus providing both the *in situ* and equilibrium values from a single set of measurements.

The perturbation of pH resulting from changes in the amount of organic acid anions may be important at lower bicarbonate ion concentrations. If during a given season there is a net production of consumption of 'organic' alkalinity, then the concentration of bicarbonate ion must change in the opposite sense to maintain the charge balance. If the P_{CO_2} remains constant through equilibration with the atmosphere, a change in the $HCO_3^-/H_2CO_3^*$ ratio produces a change in pH in accord with equation (1).

Total alkalinity (ΣAlk) has been suggested as the best measure of acidification trends in lakes and streams. Because it can be measured with a precision of better than 1%, a reliable value for the free acidity of water at equilibrium with the atmosphere could presumably be calculated using equations (1) and (2). However, organic acid anions may comprise a substantial portion of the total alkalinity^{4,12–16}. Therefore, attempts to calculate pH from measured ΣAlk without accounting for this component may be in error, especially at low $[HCO_3^-]$ ($pH \approx 6$).

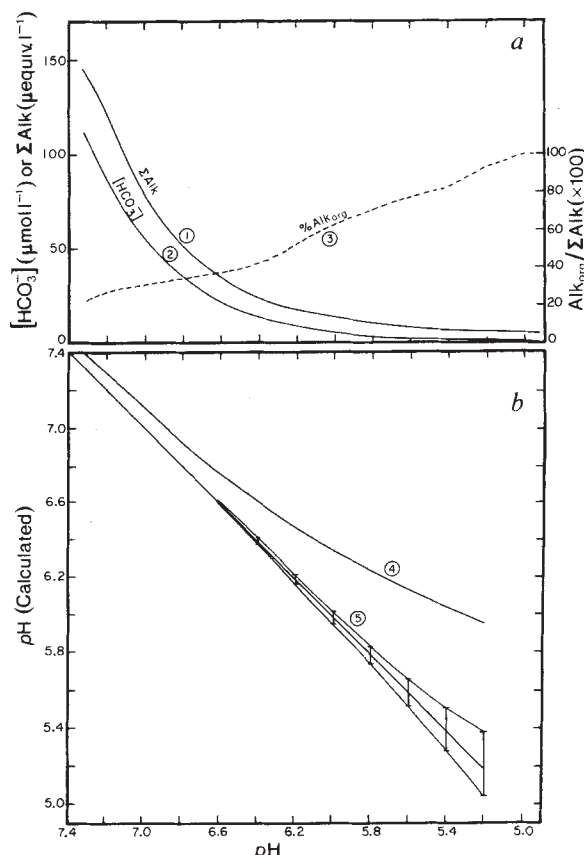


Fig. 3 *a*, The response of ELA lake 239 to acidification is shown by a simple open-system acid titration. Curves 1 and 2 represent the concentrations of ΣAlk and HCO_3^- , respectively, as a function of pH . Curve 3 indicates the fraction of ΣAlk made up of organic acid anions. $P_{\text{CO}_2} = 340 \mu\text{atm}$ at 20°C . *b*, The pH calculated for lake 239 by assuming that $\Sigma\text{Alk} = [\text{HCO}_3^-]$ is given by curve 4. These calculated pH s deviate from the true values (curve 5) by an increasing amount as the acidification progresses. Error bars on curve 5 represent the analytical uncertainty of the calculated pH values using the $P_{\text{CO}_2}/\text{DIC}$ method.

Consider again ELA lake 239, with an average $\Sigma\text{Alk} = 146 \mu\text{equiv. l}^{-1}$, $[\text{HCO}_3^-] = 113 \mu\text{mol l}^{-1}$ and, by difference, $\text{Alk}_{\text{org}} = 33 \mu\text{equiv. l}^{-1}$. If this lake were to undergo acidification, the components of alkalinity might vary with pH as shown in Fig. 3*a*. Because the pK s for organic acids are generally much lower than pK_1 for H_2CO_3^* (refs 12–16), most of $[\text{HCO}_3^-]$ will be consumed before there is a significant depletion in organic anions. At pH values near 6, ΣAlk is much greater than $[\text{HCO}_3^-]$, hence any attempt to calculate pH using K_1 will lead to serious error as is shown in Fig. 3*b* by comparing curves 4 (which assumes that $\Sigma\text{Alk} = [\text{HCO}_3^-]$) and 5 (which represents the pH calculated by the $P_{\text{CO}_2}/\text{DIC}$ method).

The method for monitoring the acidity of lakes using DIC and P_{CO_2} measurements avoids both electrode and organic alkalinity problems and minimizes perturbations caused by fluctuations in P_{CO_2} . Given the variability of our calculated pH values for lake 239 (Fig. 1), we could reliably detect a change of $>0.1 \text{ pH}$ units in the equilibrium-acid status of the lake by making a single set of midsummer $P_{\text{CO}_2}/\text{DIC}$ measurements at any time in the future. Collecting a few samples during a summer would reduce this detection limit by at least a factor of two.

The monitoring strategy described here is not intended for lakes that are already acidified. It is only applicable to waters with $\text{pH} > 5.5$, below which the difference between DIC (that is, $[\text{HCO}_3^-] + [\text{CO}_2]$) and $[\text{CO}_2]$ becomes too small to permit accurate determinations of $[\text{HCO}_3^-]$. For lakes now on the brink of serious acidification effects (that is, in the pH range 5.5–7.0), our method provides a far more reliable and sensitive indicator of acidification trends than can be obtained by any other method.

Our proposal requires only that the determination of P_{CO_2} and DIC be included in the existing monitoring programmes.

We thank T. Takahashi and R. Hesslein for valuable discussions. Some of the analyses reported here were performed by the ELA chemistry group under the direction of M. Stainton and G. Linsey. This work was supported under the auspices of NSF grant DEB 80 17639 and the Department of Fisheries and Oceans, Canada.

Received 24 October 1984; accepted 13 March 1985.

- McQuaker, N. R., Kluckner, P. D. & Sandberg, D. K. *Envir. Sci. Technol.* **17**, 431–435 (1983).
- Brezinski, D. P. *Analyst* **108**, 425–442 (1983).
- Brezinski, D. P. *Talanta* **30**, 347–354 (1983).
- Herczeg, A. L. & Hesslein, R. H. *Geochim. cosmochim. Acta* **48**, 837–845 (1984).
- Covington, A. K., Whalley, P. D. & Davison, W. *Analyst* **108**, 1528–1532 (1983).
- Tyree, S. Y. *Atmos. Envir.* **5**, 57–60 (1981).
- Schindler, D. W. *et al. Science* **177**, 1192–1195 (1972).
- Emerson, S. *Limnol. Oceanogr.* **20**, 743–753 (1975).
- Talling, J. F. *J. Ecol.* **64**, 79–122 (1976).
- Stumm, W. & Morgan, J. J. *Aquatic Chemistry. An Introduction Emphasizing Chemical Equilibria in Natural Waters* 2nd edn, Ch. 4 (Wiley, New York, 1981).
- Stainton, M. P., Capel, M. J. & Armstrong, F. A. J. *Can. Fish. Mar. Ser. Miscell. Spec. Publ.* **25**, 1–180 (1977).
- Beck, K. C., Reuter, J. H. & Perdue, E. M. *Geochim. cosmochim. Acta* **38**, 341–364 (1974).
- Brossett, C. W. *Air, Soil Pollut.* **14**, 251–265 (1980).
- Driscoll, C. T. & Bisogni, J. J. in *Modeling of Total Acid Precipitation Impacts* (ed. Schnoor, J. L.) 53–72 (Ann Arbor Science, Ann Arbor, 1984).
- Oliver, B. G., Thurman, E. M. & Malcolm, R. L. *Geochim. cosmochim. Acta* **47**, 2031–2036 (1983).
- Purdue, E. M., Reuter, J. H. & Parrish, R. S. *Geochim. cosmochim. Acta* **48**, 1257–1263 (1984).

Secular climate change in old-growth tree-line vegetation of northern Quebec

Serge Payette, Louise Filion, Line Gauthier & Yves Boutin

Centre d'études nordiques, Université Laval, Sainte-Foy, Québec, Canada G1K 7P4

The climate of the Northern Hemisphere changed during recent centuries, as shown by the Little Ice Age episode¹ and the warming trend of the past 100 years^{2,3}. The ecological impacts of these changes have yet to be evaluated in several terrestrial ecosystems, incorporating direct evidence such as detailed botanical field observations⁴. We report here results of the analysis of a ~600-year response of a lichen-spruce woodland to this long-term trend, which are thought to be the first extensive illustration of these impacts in the Subarctic. It suggests that tree-line vegetation is in a dynamic equilibrium with climate in the absence of other external disturbances; this is emphasized by spruce reaction through phenotypic adaptation—a shift from stunted individuals (krummholz) to normal trees (forest)—and differential regeneration. This study produces evidence that marginal northern forests can persist through time and that successional processes, in the absence of fire, perpetuate the original lichen-spruce facies. The longest tree-ring chronology (AD 1398–1982) yet available in eastern North America was constructed from living and dead spruces found in the lichen woodland of our study.

It has been shown recently⁵ that the extensive treeless areas within the northern Québec Forest-Tundra zone are the result of late Holocene fire disturbance and climatic change. Exceptions to this deforestation trend are found around the tree line near Hudson Bay, where old-growth coniferous vegetation persisted without any fire for several centuries. One of the best examples is a lichen-black-spruce (*Picea mariana* (Mill.) BSP.) woodland, 3–4 km long, about 10 km south of the present tree line in the Bush Lake area ($57^\circ 47' \text{N}$, $75^\circ 45' \text{W}$). This stand occupies a well-drained site that burned for the last time around 900–850 yr BP, in the immediate vicinity of the sampling area, as indicated by two radiocarbon-dated charcoals found below the organic layer (UQ-163: 890 ± 55 yr BP; UQ-158: 850 ± 85 yr BP), that is, around dendroyears AD 1075 ± 57 and 1160 ± 86 using the radiocarbon-dated calibration curve of Stuiver⁶.

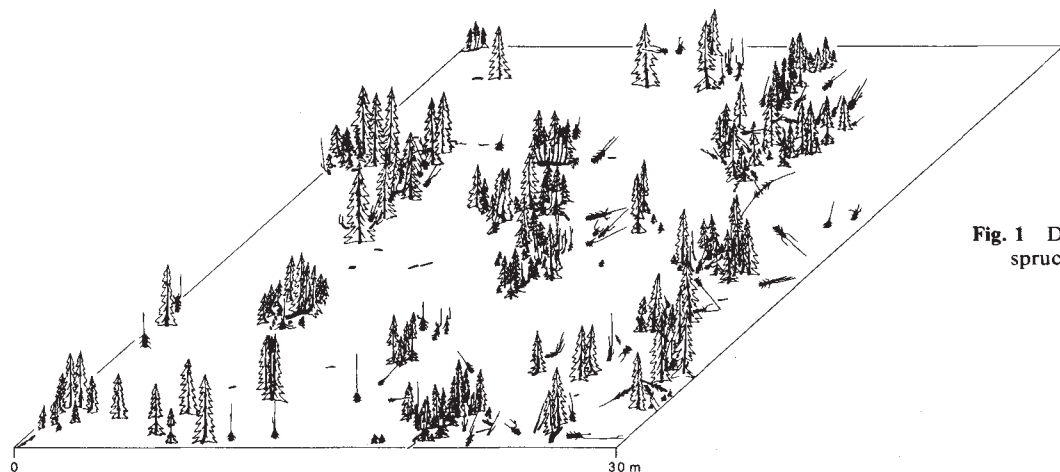


Fig. 1 Distribution of the living and dead spruces within the 900 m² quadrat.

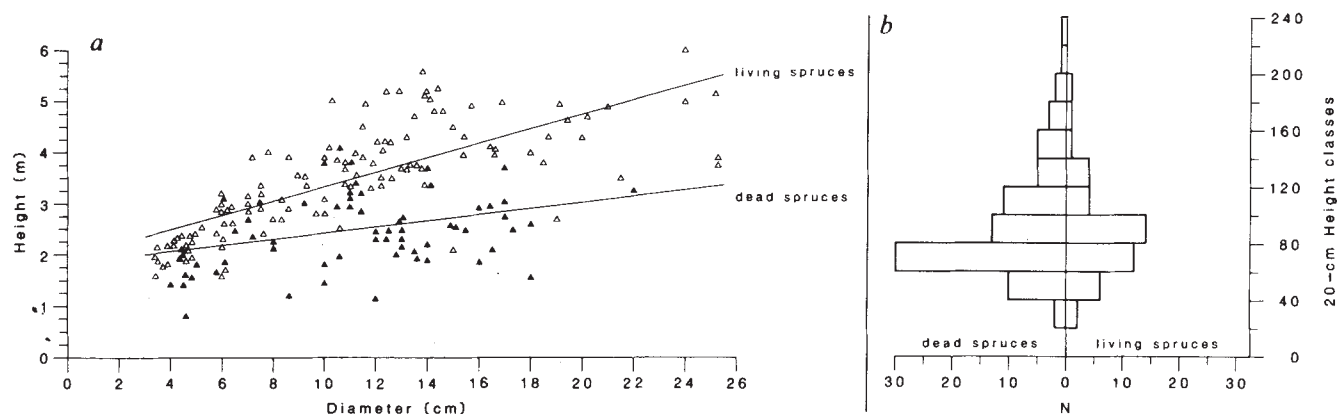


Fig. 2 *a*, Relationship between basal stem diameter and total height in living (Δ , $y = 1.9621 + 0.1391x$, $r = 0.74$) and dead ($\blacktriangle$, $y = 1.8203 + 0.0604x$, $r = 0.35$) spruces. Differences in the regression slopes are statistically significant at the 0.001 level. *b*, Height (20-cm class) above ground of multiple-stem levels and branch leaders. The bulk of multiple-stem level distribution in dead and living spruces is respectively around 60–80 and 80–100-cm classes, which suggests an overall difference of 20 cm in depth of snow cover; the dead spruces were probably in an environment with less snow compared with living spruces.

The woodland consists of dominant living low spruce trees along with shrubby or subarborescent stunted leaning dead spruces (Fig. 1). We have postulated that the two contrasting growth forms developed under different climatic regimes; the abundance of dead spruces in such an environment made us suspect that the woodland was growing old in the absence of external disturbance—a unique situation for evaluating the stability of marginal northern forests.

A 900 m² (30m × 30m) quadrat was used to determine the spatio-temporal development of the coniferous vegetation (Fig. 1). All the living and dead spruces were mapped and their growth forms, such as arboreal, subarborescent and shrubby, described (Fig. 1). Diameter at stem base, total height and winter damage expressed by morphological traits such as above-ground multiple-stem and branch leaders, along with bare (leafless and branchless) stems above snow cover, were measured on all individuals (Fig. 2). The origin, by layering or by seed, was determined from stem morphology, stem distribution and below-ground branch connections⁷. Samples of living spruces were taken at stem base and at the multi-stem level that is thought to mark the mean winter snow position. The dead stems were sampled immediately above the level of decomposition (stems rotted near the ground), generally 30–100 cm above ground level, and also at the snow–air interface and stem apex; although the real age, that is, the age of the stem at the ground level, could not be determined on these samples, the data were used to estimate the date of stem death.

Tree-ring measurements were performed on fine-sanded cross-sections with a standard micrometer (0.01 mm precision). The

dating of dead stems was made possible because of the very good condition of preservation of these samples and the presence of such diagnostic light rings of very few latewood cells as those dated to AD 1593, 1816 and 1817 (L.F., S.P. and L.G., unpublished results). The tree-ring chronology was extended to AD 1398 by cross-dating. The oldest black spruce was 504 yr old, which extends the known longevity for that species⁸ by at least 200 yr. The chronology was built from 39 radii belonging to 26 stems. Reaction wood⁹ was common in leaning stems, but measurements made on the different sets of tree-rings, with or without compression wood, showed the same growth; samples with a changing direction of compression wood were rejected. Missing or incomplete rings were frequent, especially for the period 1600–1610, and annual rings were particularly narrow between 1398 and 1880, virtually at the lower limit of spruce growth. Because of these anatomical and morphological ring characteristics, the chronology has been based on the arithmetical mean of the 39 series. In addition, an overall-growth curve was produced from multiple polynomial fitting of the 39 series.

The 233 stems recorded consist of 136 living and 97 dead specimens (Fig. 1). Striking size differences were found between the two groups, especially with increasing size of basal stem diameter (Fig. 2*a*); large dead stems reached about 3.5 m in height, whereas large living spruces were 1–1.5 m higher. Living spruce boles are predominantly erect, radially symmetrical and, occasionally, eroded at the snow–air interface level, whereas dead spruces showed oblique and anisomorphic boles together with multiple-stem leaders located at a former snow–air interface

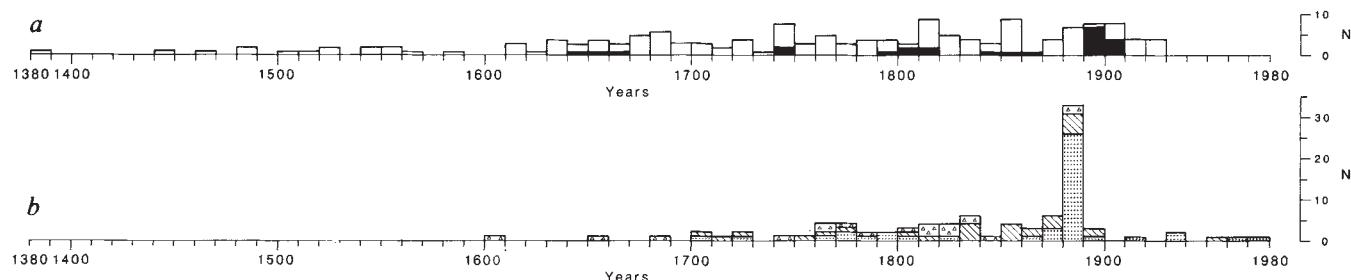


Fig. 3 *a*, Age structure of living spruces according to origin (white: layering; black: seed). Age distribution among 10-yr classes is thought to cover the difference between real stem age and stem age at the sampling level. Because the age structure is of the present-day (sampling year in 1983) living individuals, the data cannot reliably be extrapolated back in time. Absence of seedling establishment in the climatically favourable period of the 1930s and 1940s appears to be related to quadrat sampling size; areally small clustered cohorts from that period were found a few metres away, but are not incorporated here in the age structure histogram. We think that the naturally low seedling regeneration rate of this lichen woodland during the past centuries could be estimated more properly with larger sampling units. *b*, Mortality dates of spruces (triangle, death of stem apex; hatched, death of stem at the snow-air interface; dot, death of basal stem, that is, below the snow-air interface). 10-yr class. Mortality dates were estimated only for well-preserved outer stem sections, which give the overall mortality trend of established stems during the past two centuries, but not for previous centuries because of an unknown proportion of decomposed stems.

level (Fig. 1). Frequency distribution of postulated snow-air levels between living and dead spruces suggests that the former were growing in an environment with less snow (Fig. 2*b*), under colder conditions probably associated with frost desiccation and deep-freezing of exposed leaves^{10,11}.

The all-aged structure of the living spruce population (Fig. 3*a*) indicates continuous development during the past centuries, mainly through layering. Only 16% were identified as seed-origin individuals and 50% of these germinated between 1890 and 1910; the remaining seed population established episodically between 1640 and 1870. Layering occurred continuously throughout the past 600 yr (young layers were not sampled within the quadrat). Mortality was also an important component of the spruce population dynamics, particularly between 1700 and 1980 (Fig. 3*b*). Tree-ring analysis of dead stems revealed differential mortality according to the position of the living tissues relative to snow cover. Above-snow stems often died 100 yr earlier than stem sections located at the snow-air interface; death of the trunk followed within a few decades. An important feature here is the rather sudden mass mortality of spruces during a 10-yr period, between 1880 and 1890, when 57% of the subfossil population died. At that time, 70% of the dying spruces were older than 200 yr and 30% were older than 250 yr. Senescence appears to be one factor explaining the mass mortality; considerable reduction of the photosynthetic structure is a possible factor, as suggested by the very low growth during many decades before death. The cost of reproduction¹² may have also some influence here on the overall vigour of the old-aged and moribund spruces at a time when cone production was stimulated. In fact, this period of mass mortality coincides with a short interval of seedling establishment (Fig. 3*a, b*). Other factors such as pathogen attacks (for example, fungus, insects) remain uncertain.

Tree-ring responses to the long-term developmental process of spruce vegetation are clearly expressed in the 600-yr chronology (Fig. 4). Two periods of different duration are shown, a slow-growth period between 1398 and 1890 and a faster growth period afterwards. Very narrow rings were formed during the first period, especially around 1420–1435, 1480–1495, 1600–1610, 1690–1700, 1825–1830, 1835–1840 and 1855–1860. Most of the period corresponds to the Little Ice Age, particularly between 1550 and 1850 (ref. 1). Yearly variations in ring width were not important except before 1520; but, more data are needed to evaluate thoroughly the period before 1550 where short intervals of larger tree rings and presence of tree-like subfossils suggest less severe climates than during the Little Ice Age. Low stature and limited photosynthetic mass during that period could have caused variability in tree-ring responses, although the overall ring pattern is one of climate suppression on tree growth. Large tree-ring widths during the second period

are associated with milder climates well documented by instrumental records². Yearly variability is higher here and tree growth climaxed in the 1930s and early 1940s, during maximum warmth². A reversal in the growth trend occurred afterwards and corresponds to the post-1940s cooler conditions registered in the Northern Hemisphere, except for a brief interval between 1950 and 1960 (ref. 14).

The Bush Lake lichen woodland represents an outstanding example of vegetation response to climate change in the absence of other known external disturbances. The discovery of older dead spruces may extend the chronology to 800–900 yr. The ecological impacts of the Little Ice Age climate were significant in terms of spruce morphology and regeneration, because the population formed a krummholz (stunted spruces) of predominant long-lived layers. However, episodic seedling establishment occurred during short intervals, such as 1640–70, 1740–50, 1800–20 and 1840–60. An important conclusion here is that krummholz may regenerate sexually during short milder periods, as suggested by the tree-ring chronology. Another is that shifting from krummholz to forest expresses a major change towards milder conditions culminating during this century. This shifting in growth habit appears to be a large-scale phenomenon observed throughout the northern Québec Forest-Tundra. It is also important to consider that the time lag in growth-form shifting, attributable to change in climate, is determined in part by the time when the critical position of apical buds and needles relative to the winter snow pack is reached. This suggests that winter climate defines the climate threshold where apices of growing spruces, or exposed buds, needles and twigs, located at the snow-air interface, experience harsh conditions that eventually discriminate the future growth forms among long-lived individuals. Seedlings established during the past two centuries that have developed into tree form, probably reached the snow-air interface when the winters became somewhat milder. Therefore, we may conclude that the time lag between a change in

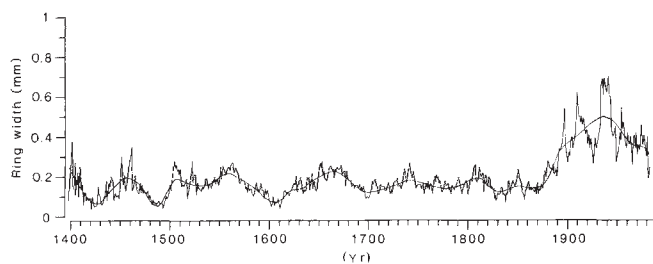


Fig. 4 Tree-ring chronology, AD 1398–1982, built from 26 stems of living and dead spruces (mean ring width in mm from 39 radii). Also included is the long-term growth curve derived from polynomial fitting¹³.

climate and phenotypical spruce response is many years, sometimes a few decades, a situation controlled in growing spruces by apical-bud residence time into the critical snow-air interface. The transition from the Little Ice Age to the present regime was also characterized by mass mortality in old-aged spruces that were not able to take advantage of the new favourable climates and by the replacement of these spruces by seedlings that later formed normal trees. Although we know of no other similar example of mass mortality in northern forests that is not caused apparently by insect outbreaks, we believe that our results suggest a possible link with the changing ecological conditions that occurred at the end of the nineteenth century, when old spruces with limited photosynthetic mass were already engaged in a lengthy process of dying. A closer look at the biological effects of energy expenditure in senescent or moribund spruces during a period of climate recovery, could lead to a better understanding of this intriguing phenomenon. Finally, we conclude that the overall developmental sequence described here indicates that lichen woodlands can regenerate sexually in the absence of fire disturbance and do not degrade into tundra vegetation as suggested elsewhere¹⁵. It illustrates the importance of vegetation inertia, in a fire-free ambience, for persistence of communities in cold environments characterized by episodic sexual regeneration. From a climatological viewpoint, the conditions expressed by this old-growth vegetation were probably associated during the Little Ice Age with very cold windy winters with thin snow cover and cool moist summers and, during the recent warming, with milder and snowier winters (a postulated 20-cm mean

increase in snow depth based on growth forms, Fig. 2b) and mild summers. The 600-yr chronology exhibits the marked differences observed among dead and living spruces and it may be considered as a typical example of secular climatic change in the eastern part of the North American Subarctic.

We thank Professor J. C. Ritchie (Toronto) for revising our manuscript and for useful suggestions. This study has been supported financially by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Ministère de l'éducation of Québec (FCAC Program). We also thank Mr. Ivan Grenier for drafting the figures, Dr Paul Comtois for field assistance.

Received 28 December 1984; accepted 7 March 1985.

1. Lamb, H. H. *Climate, Present, Past and Future* Vol. 2 (Methuen, London, 1977).
2. Jones, P. D., Wigley, T. M. L. & Kelly, P. M. *Mon. Weath. Rev.* **110**, 59-70 (1982).
3. Kelly, P. M., Jones, P. D., Sear, C. B., Cherry, B. S. G. & Tavakol, R. K. *Mon. Weath. Rev.* **110**, 71-83 (1982).
4. Oliver, C. D. & Stephens, E. P. *Ecology* **58**, 562-572 (1977).
5. Payette, S. & Gagnon, R. *Nature* **313**, 570-572 (1985).
6. Stuiver, M. *Radiocarbon* **24**, 1-26 (1982).
7. Lègère, A. & Payette, S. *Arct. Alp. Res.* **13**, 261-276 (1981).
8. Fowells, H. A. *Silvics of Forest Trees of the United States* (Agricultural Handbook no. 271, U.S.D.A., 1965).
9. Scurfield, G. *Science* **179**, 647-655 (1973).
10. Tranquillini, W. *Physiological Ecology of the Alpine Timberline* (Springer, Berlin, 1979).
11. Wardle, P. *Arct. Alp. Res.* **13**, 419-423 (1981).
12. Silvertown, J. W. *Introduction to Plant Population Ecology* (Longman, London, 1982).
13. Fritts, H. C. *Tree Rings and Climate* (Academic, New York, 1976).
14. Yamamoto, R. *Proc. Tech. Conf. Climate*, 360-375 (World Climate Programme, World Meteorological Organization no. 578, 1980).
15. Strang, R. M. *Can. J. For Res.* **3**, 140-143 (1973).

Isotope assessment of Holocene human diets in the southwestern Cape, South Africa

Judith C. Sealy & Nikolaas J. van der Merwe

Department of Archaeology, University of Cape Town, Private Bag, Rondebosch, 7700 Cape Town, South Africa

Models of seasonal mobility to exploit seasonally abundant food sources have been proposed for prehistoric hunter-gatherers in many parts of the world¹⁻³. Some such hypotheses involve fundamental and insufficiently tested assumptions about the nature of both hunter-gatherer societies and the archaeological evidence that they leave. The present study is an independent test of such a hypothesis proposed for the southwestern Cape of South Africa. In this strongly ecologically differentiated area there are four distinct ecological zones that would have offered four different sets of resources to prehistoric people. Obvious modern seasonal fluctuations in these resources, plus a considerable amount of archaeological evidence, led to the suggestion that prehistoric hunter-gatherers moved in a regular seasonal cycle across the zones^{4,5}; this would have allowed them to make maximum use of temporarily plentiful plant and animal foods in some areas, while avoiding lean periods in others. However, as reported here, direct measurements of food intake, as reflected in the stable carbon isotope ratios of archaeological human skeletons, reveal that this was not the case. The implications of this study extend beyond the relevance to local archaeology to more general questioning of the ways in which archaeological data should be used to generate hypotheses.

Figure 1 shows the four north-south ecological zones of the southwestern Cape: the Atlantic shoreline, the temperate coastal plain, the Fold Mountain Belt and the semi-desert Karoo. Each zone has a different geology, rainfall and vegetation. Holocene archaeological sites in the area contain quantities of plant and animal food waste, much of it identifiable as having been collected in the summer (in the mountains) or in the winter (on the coast).

Inland sites are rich in the remains of edible plants: corm casings from starchy Iridaceae corms (at their most palatable

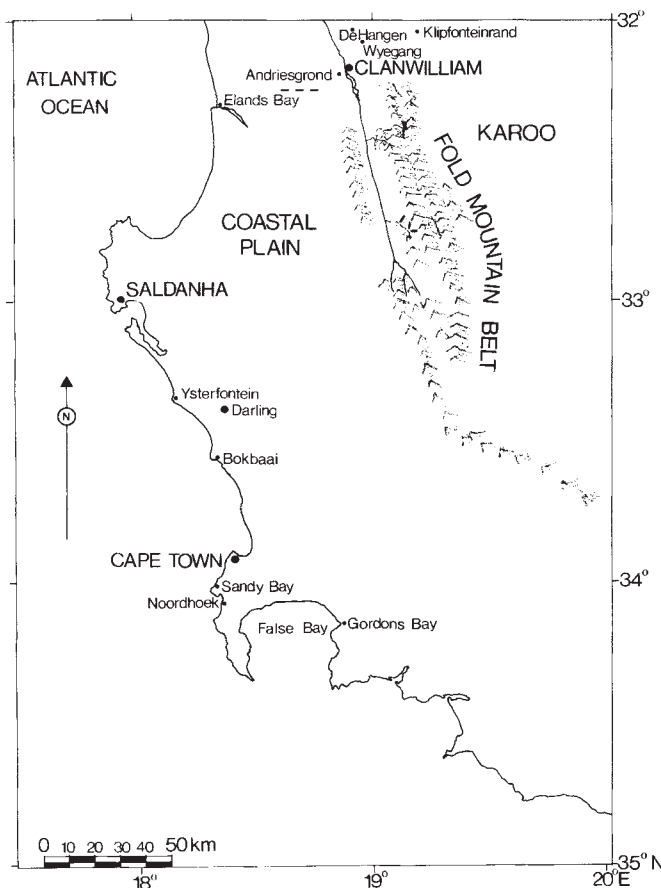


Fig. 1 Map showing geographical features of the southwestern Cape and localities of the skeletons listed in Table 1.

in late summer) and seeds from the fruits of summer-fruited plants. The faunal assemblages are dominated by tortoise bone (tortoises are most active, and hence easiest to collect, in summer). The second most common animal is the rock hyrax (*Procavia capensis*). Examination of the stage of tooth eruption in juvenile hyrax mandibles reveals that almost half of those from

Table 1 $\delta^{13}\text{C}$ values of dated archaeological human skeletons

Archaeometry lab. no.	Skeletal register no.	Locality	Date (yr BP)	$\delta^{13}\text{C}$ (‰)	
Coastal					
UCT 1050	SAM-AP 1153	False Bay	Associated pottery (?)	-15.1	<2,000 BP
UCT 622	UCT 158	Sandy Bay	Associated pottery (?)	-15.0	
UCT 623	UCT 159	Sandy Bay	Associated pottery (?)	-13.6	
UCT 625	UCT 248	Noordhoek	Associated pottery (?)	-12.9	
UCT 459	SAM-AP 5083	Ysterfontein	Pta-926 1,490 ± 55	-14.5	
UCT 1048	UCT 60	Saldanha	Pta-2,005 955 ± 50	-14.6	2,000-4,000 yr BP
UCT 451	SAM-AP 4813	Bokbaai	Associated pottery (?)	-15.1	
UCT 1049	SAM-AP 1443	Gordon's Bay	Pta-2,309 2,050 ± 50	-11.8	
UCT 1052	UCT 162	Ysterfontein	Pta-929 2,880 ± 50	-11.5	
UCT 1092	-	Eland's Bay	Pta-1,754 3,835 ± 50	-14.7	
UCT 1051	UCT 112	Darling Coast	Pta-2,003 4,445 ± 50	-11.2	4,000-8,000 yr BP >8,000 BP
UCT 1091	-	Eland's Bay	Pta-1,829 8,000 ± 95	-14.6	
UCT 206	-	Eland's Bay	Pta-3,086 9,750 ± 100	-12.3	
UCT 202	-	Eland's Bay	OxA-478 10,860 ± 180	-12.5	
Inland					
UCT 457	SAM-AP 1449	Clanwilliam	OxA-453 2,230 ± 100	-17.3	2,000-4,000 yr BP
UCT 203	UCT 334	Andriesgrond	OxA-457 3,850 ± 80	-20.2	
UCT 1093	-	Klipfonteinrand	Pta-1,642 3,540 ± 60	-19.6	4,000-8,000 yr BP
UCT 1053	UCT 331	Wyegang	Pta-1,649 6,900 ± 70	-18.4	

All pottery so far found in the area post-dates 2,000 yr BP; its association with skeletons therefore constitutes an approximate guide to their ages. Some of these skeletons (those from Saldanha South) lie outside the area for which the seasonal mobility hypothesis was originally formulated. They are included mainly to demonstrate the consistently positive character of coastal isotope values. Pta, radiocarbon dates from the Pretoria laboratory; OxA, dates from the Oxford radiocarbon accelerator.

the cave of De Hangen (the site with the largest faunal assemblage) were probably killed between October and May. Finally, historical accounts of meetings between European travellers and 'bushman' hunter-gatherers in the area all refer to summer encounters.

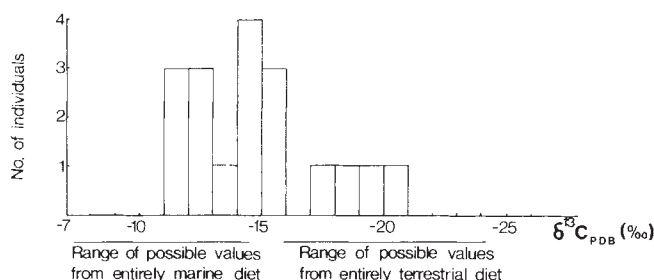
The coastal segment of the seasonal round is represented at sites such as Elands Bay Cave: like most coastal sites, this has few plant food remains but large quantities of limpet and mussel shells. If shellfish were a common, reliable resource taking the place of plant foods as a coastal staple⁴⁻⁶, Parkington⁴ suggested that it would have been safer to exploit them in winter. Winter outbreaks of the poisonous dinoflagellate *Gonyaulax catenella* are less common than summer ones; these diatoms render filter-feeding shellfish (such as mussels) toxic to man. The most common marine mammal from the Elands Bay deposits is the Cape Fur Seal (*Arctocephalus pusillus*). Young seals are born in November and nearly all the excavated mandibles are those of yearlings that died between July and October; seasonal rings on a small sample of canines from adult male seals obtained from the same site apparently also reflect this short time period⁷. The bones of several species of birds that are winter visitors to the nearby Verloren Vlei estuary (for example, flamingos) are found throughout the Holocene levels of the cave deposits. Winter is also the time when the Vlei, swollen by winter rains, used to break through into the sea, carrying with it organic material that would have attracted fish such as the white steenbras (*Lithognathus lithognathus*), common in the cave deposits.

Thus, there is a considerable amount of archaeological evidence for summer occupation at De Hangen and winter occupation at Elands Bay Cave. Inland and coastal sites have been seen as different segments of one integrated subsistence economy. People supposedly moved between the various resource zones in such a way as to exploit peaks in animal and, particularly, plant resources, while avoiding troughs. Recently, Parkington⁵ has fitted his seasonal mobility hypothesis more securely into a chronological framework. He now expects that it will be found to apply, in the form outlined above, mainly in the period between 4,000 and 2,000 yr BP; before this, from 8,000 to 4,000 yr BP, there was a hiatus in the occupation of Elands Bay Cave and of at least one other site nearby, while after 2,000 yr BP the system was probably seriously disrupted by the emergence of pastoralists with domesticated animals on the coastal plain.

The seasonal mobility hypothesis is directly testable by measuring the $^{13}\text{C}/^{12}\text{C}$ ratio. The use of $\delta^{13}\text{C}$ measurements ($\delta^{13}\text{C}_{\text{sample}} = [^{13}\text{C}/^{12}\text{C}_{\text{sample}} / ^{13}\text{C}/^{12}\text{C}_{\text{standard}}] - 1 \times 1,000\text{‰}$) as a source of dietary information is now an established archaeological technique^{8,9}. Most previous studies incorporating such measurements have, however, been limited by having to use global averages for the isotope ratios of the dietary base. In the study reported here, we began by analysing samples of all the foods that we know were important in the diet of prehistoric humans in the southwestern Cape. We also included a cross-section of other edible organisms that would leave no trace in the archaeological record.

The results of the analyses form very clear patterns, which are presented in full elsewhere¹⁰. In summary, the mean $\delta^{13}\text{C}$ value of the marine-based foods consumed by people living on the coast (shellfish, crayfish, fish, seal and seabird meat) is $-15.6 \pm 1.6\text{‰}$ (range = -12.3 to -19.4‰ , $n = 29$). The edible parts of terrestrial plants have a mean $\delta^{13}\text{C}$ of $-25.4 \pm 1.8\text{‰}$ (range = -22.3 to -29.2‰ , $n = 26$) whereas terrestrial animal meat has a mean of $-23.6 \pm 1.3\text{‰}$ (range = -20.9 to -25.8‰ , $n = 15$). Therefore, a marine-based diet has an isotope 'signature' substantially different from a terrestrially based one; this difference would be reflected in the bones of people consuming these foods (the absolute values for bones being some 5‰ more positive as isotope discrimination proceeds stepwise along trophic levels).

Next, we analysed several archaeological human skeletons, some of which were found at the coast and some in the mountains. If they represent a seasonally transhumant population they should all have broadly similar $\delta^{13}\text{C}$ values intermediate

**Fig. 2** Histogram of the results given in Table 1.

between the values for purely marine and purely terrestrial diets; it is obvious from Table 1 that this is not the case. Instead, the coastal skeletons reflect a predominantly marine diet and the inland skeletons an almost entirely terrestrial diet. Clearly, whatever seasonal round either group might have engaged in, they do not represent segments of the same subsistence economy.

The coastal skeletons show $\delta^{13}\text{C}$ values that range from almost entirely marine-derived (-11%) to values which could reflect up to 50% terrestrial food composition (-15%). We consider that the sample is as yet too small to detect geographical or chronological patterning, though this will be one of the principal aims of further work. The most interesting point is not that some coastal people consumed terrestrial foods (available, in any case, near the shore) but that, in direct contrast to nearly all models of prehistoric coastal subsistence^{4-6,11}, a significant proportion of the coastal population consumed an entirely marine diet and so must have spent all its time on the shore. Seventeenth-century European settlers described 'strandlopers' doing this¹²; from the isotope data it seems to be a pattern of life extending back at least as far as the beginning of the Holocene.

This result may seem surprising in view of the extremely plausible arguments for seasonal mobility summarized above. Closer examination of these arguments, however, reveals that although the evidence is consistent with such a hypothesis, it is also extremely inconclusive¹⁰. Although people certainly were living in the mountains in the summer and on the coast in winter, the hypothesis rests on the unproven assumption that they were the same people and that they were not living in the mountains in winter, or on the coast in summer—this would be much more difficult to demonstrate, but it is essential to do so to postulate

population movement between the two areas. We suggest that comparable hypotheses formulated for other parts of the world might benefit from similar critical re-examination.

We thank John Parkington, who suggested this project, provided the archaeological samples, and partly financed the project through his Spatial Archaeology Research Unit, supported by the Human Sciences Research Council of South Africa. Other colleagues also gave us free access to ideas and unpublished data, and we thank the many people who helped to collect samples. We thank John Vogel of the CSIR in Pretoria and the Radiocarbon Accelerator Laboratory at Oxford University for the radiocarbon dates. Additional financial support was provided by the Council for Scientific and Industrial Research and the Harry Oppenheimer Institute for African Studies at the University of Cape Town.

Received 4 December 1984; accepted 7 February 1985.

1. Higgs, E. S., Vita-Finzi, C., Harris, D. R. & Fagg, A. E. *Proc. prehist. Soc.* **33**, 1-29 (1967).
2. Higham, C. in *Problems in Economic and Social Archaeology* (eds Sieveking, G., Longworth, I. H. & Wilson, K. E.) 221-233 (Duckworth, London, 1976).
3. Schrire, C. in *Models in Archaeology* (ed. Clarke, D. L.) 653-670 (Methuen, London, 1972).
4. Parkington, J. E. *Afr. Stud.* **31**, 223-243 (1972).
5. Parkington, J. E. in *Advances in World Archaeology* (eds Wendorf, F. & Close, A.) 89-142 (Academic, London, 1984).
6. Deacon, H. J. S. *Afr. Archaeol. Soc. Goodwin Ser.* **1**, 26-45 (1972).
7. Fletemeyer, J. R. S. *Afr. archaeol. Bull.* **32**, 146-149 (1977).
8. Vogel, J. C. & van der Merwe, N. J. *Am. Antiq.* **42**, 238-242 (1977).
9. van der Merwe, N. J. *Am. Scient.* **70**, 596-606 (1982).
10. Sealy, J. thesis, Univ. Cape Town (1984).
11. Deacon, H. J. S. *Afr. archaeol. Soc. Monogr. Ser.* **1** (1976).
12. Thom, H. B. (ed.) *Journal of Jan van Riebeeck* Vols 1, 2, 3, (A.A. Balkema, Cape Town, 1952, 1954, 1958).

Evidence for echolocation in the oldest known bats

Michael J. Novacek

Department of Vertebrate Paleontology, American Museum of Natural History, New York, New York 10024, USA

The earliest-known bats are represented by excellent fossil material, including virtually complete skeletons of *Icaronycteris index* from the early Eocene (50 Myr BP) of western Wyoming^{1,2} and *Palaeochiropteryx tupaiondon* from the middle Eocene (45 Myr BP) 'Grube Messel' of western Germany^{3,4}. These taxa have been closely allied^{1,5} with Recent Microchiroptera, a suborder of diverse bats noted for their powers of ultrasonic echolocation⁶. A problem with this relationship is the alleged absence in the Eocene forms of specializations in the auditory region⁷ and other aspects of the skeletal system^{1,2}. It has been proposed, therefore, that the oldest bats are members of a group more primitive and possibly ancestral to the Microchiroptera and the visually oriented Megachiroptera⁸. Previously undescribed specimens now show, however, that *Icaronycteris* (Fig. 1) and *Palaeochiropteryx* (Fig. 2) share special basicranial features with microchiropterans which suggest comparable refinement of ultrasonic echolocation. These results support the theory that a sophisticated sonar system was present in the earliest records of microchiropteran history.

The ultrasonic echolocation of living microchiropterans has been related to the structure of the middle-ear ossicles⁹⁻¹⁵, laryngeal-hyoid apparatus^{16,17} and the cochlea^{9,14,18-20}. The ear ossicles are not well preserved in *Palaeochiropteryx*, but in *Icaronycteris* there is a partial malleus with a large orbicular apophysis (Fig. 1). This rounded process is well developed in microchiropterans, but is absent or weak in non-echolocating megachiropterans⁹⁻¹³. In microchiropterans the ossicles are tightly articulated for optimal tuning at ultra-high frequencies¹²⁻¹⁴. Mechanical analysis¹³ suggests that the orbicular apophysis shifts the centre of mass of the malleus away from the rotational axis of the ossicle chain, increases the moment

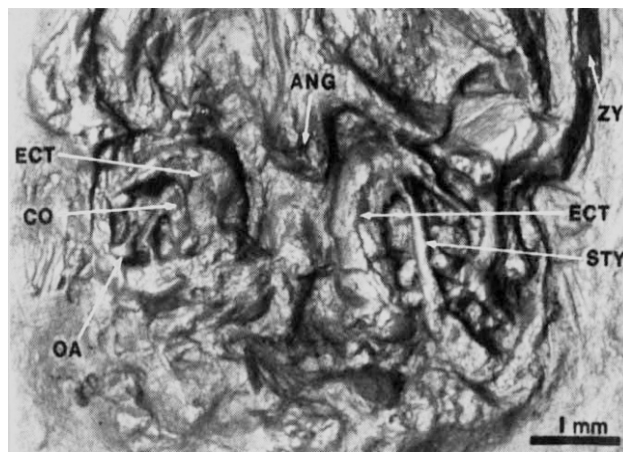


Fig. 1 Ventral aspect of cast of posterior skull of *Icaronycteris index*, University of Wyoming (UW) 2244. ANG, angular process of lower jaw; ECT, ectotympanic ring; CO, cochlea; OA, orbicular apophysis of malleus; STY, stylohyal; ZY, zygomatic arch. In this specimen the stylohyal lies lateral to its normal position ventral to the ectotympanic.

of inertia and lowers the natural frequency of the ossicle system. The apophysis may therefore contribute to the attenuation of natural frequencies that might otherwise be too high, even for echolocating bats¹³.

Icaronycteris (Fig. 1) and *Palaeochiropteryx* (Fig. 2) both have a well-developed stylohyal element. In megachiropterans, as in most mammals, the stylohyal is usually a weak filamentous structure easily lost in macerated skulls²¹. The stylohyal in microchiropterans is the site of well-developed muscles that constrict the pharynx and brace the hyoid apparatus^{16,17}. The hyoid stabilizers allow the extrinsic laryngeal muscles to pull from the hyoid and move the larynx forward¹⁷, an action that precedes the emission of a series of high-frequency pulses²². The stylohyal, then, is an important component of a laryngeal-hyoid apparatus, whose variable design influences the charac-

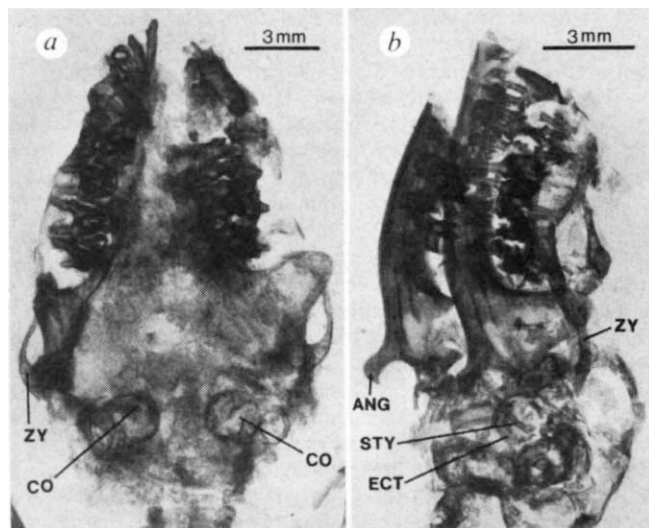


Fig. 2 *Palaeochiropteryx tupaodon*. Positive prints of radiographs of a dorso-ventral view of Senckenberg Naturmuseum und Forschungsinstitut SMF-ME 788-B (a) and a lateral view of SMF-ME 1127 (b). Abbreviations as in Fig. 1.

teristics of high-frequency vocalizations in different bat species¹⁷.

The most striking resemblance between these Eocene bats and living microchiropterans is in the large relative size of the cochlea. The dimensions of this structure are easily determined in *Palaeochiropteryx* (Fig. 2). In *Icaronycteris*, the cochlea is less well preserved, but an estimate of its size can be made from measurements between auditory structures adjacent to the cochlea (Fig. 3). The maximum width of the cochlea is a meaningful functional parameter because it is equal to the maximum width of the basal turn, that region of the cochlea especially receptive to high-frequency vibrations in adult mammals^{14,19,20,23-25}. A plot of maximum cochlear width against skull length for 96 species of bats shows no overlap between Megachiroptera and Microchiroptera, and supports the observation^{14,15,18} that the cochlea in microchiropterans is unusually large relative to the size of the skull. The great expansion and increased torsion of the cochlea in these bats is accommodated by a unique pattern of arrested ossification of the petrosal and adjacent elements during ontogeny²⁶. *Palaeochiropteryx* and *Icaronycteris* fall well within the range of relative cochlear size for Microchiroptera and exceed values for several living species within this suborder (Fig. 3). Thus, the implication is strong that these early bats share with living echolocating bats a cochlea of comparable ontogeny and function.

The basicranial evidence reviewed here supports the contention^{1,5} that the oldest bats share several derived features with microchiropterans and are members of that group. Nevertheless, the fossil bats, particularly *Icaronycteris*, retain a number of very primitive features in the sternum and distal appendages^{1,2,8}. This mosaic of traits suggests that the development of a sophisticated auditory and laryngeal system preceded the acquisition of certain other structures. Ultrasonic echolocation and its correlative morphology have been demonstrated in all living microchiropterans so far studied^{14,15,26-28}. Thus, evidence from living taxa and the fossil record are consistent with the argument that ultrasonic echolocation is a primitive sensory mode for microchiropterans.

Echolocation also occurs in the megachiropteran genus *Rousettus* and in certain cetaceans, seals, insectivorans, rodents, birds and even some blind humans²⁸. Microchiropterans are unique, however, in producing echolocation calls that are structured in frequency changes over time, a mode particularly effective in hunting volant, nocturnal prey. Other echolocators, including *Rousettus*, use transient, broad-band clicks²⁸.

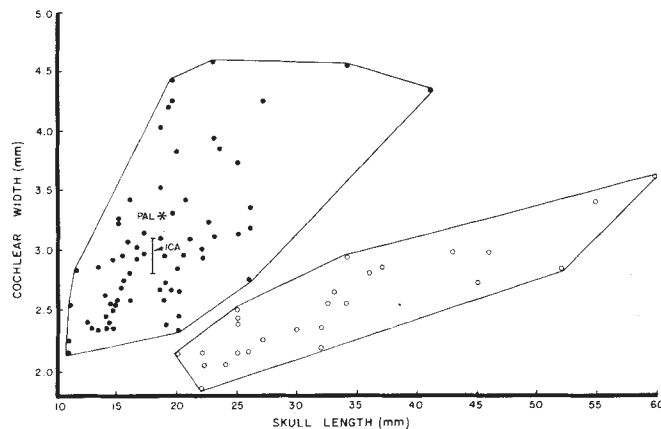


Fig. 3 Plot of maximum width of cochlea (ordinate) against skull length (abscissa) for 96 species of Microchiroptera (●), Megachiroptera (○), *Palaeochiropteryx* (*) and *Icaronycteris* (vertical line represents range of error in estimate of maximum cochlear width). Points represent 17 of 19 Recent bat families (two monotypic families of microchiropterans, the Myzopodidae and Craseonycteridae, are omitted). Convex polygons define the outer boundaries of values for each of the two suborders. Cochlear width has a weaker correlation and more positive allometry with skull length in microchiropterans ($r = 0.64$, slope = 0.80) than in megachiropterans ($r = 0.92$, slope = 0.30).

Moreover, no other mammalian group shows an auditory architecture and ontogeny closely similar to that of microchiropterans²⁶. It seems unlikely, then, that ultrasonic, time-structured echolocation was an early or primitive condition for all bats, only to be secondarily lost in the visually oriented megachiropterans. A more plausible notion is that ultrasonic echolocation was derived within bats, possibly from a condition where low-intensity, broad-band echolocation was used primarily for communication and gathering of habitat information, while vision, olfaction and direct audition (listening to sounds of prey) were used in the search for food²⁸. This transition, according to the evidence reported here, occurred earlier than 50 Myr ago, the time of first appearance of bats in the fossil record.

I thank J. Franzen (Senckenberg Museum) and J. A. Lillegren (University of Wyoming) for loan of the fossils and M. C. McKenna, A. Wyss, M. Norell, K. Koopman, L. Flynn and J. Flynn for comments. H. Tuengerthal, R. Gooris, L. Meeker and C. Tarka prepared the figures. Research was supported by the Frick Laboratory Endowment Fund (American Museum of Natural History).

Received 24 December 1984; accepted 20 February 1985.

- Jepsen, G. L. *Science* **154**, 1333-1339 (1966).
- Jepsen, G. L. in *Biology of Bats* Vol. 1 (ed. Wimsatt, W. A.) 1-64 (Academic, New York, 1970).
- Revilliod, P. *Abh. hess. geol. Landesanst.* **7**, 157-201 (1917).
- Richter, G. & Storch, G. *Natur. Mus., Frankf.* **110**, 353-367 (1980).
- Smith, J. D. & Storch, G. *Senckenberg. biol.* **61**, 153-167 (1981).
- Koopman, K. F. in *Orders and Families of Recent Mammals of the World* (eds Anderson, S. & Jones, J. K.) 145-186 (Wiley, New York, 1984).
- Segal, W. *Fieldiana, Zool.* **58**, 103-108 (1971).
- Van Valen, L. *Evol. Theory* **4**, 103-121 (1979).
- Staneke, V. J. *Nakladen Ceske Akad. Ved a Umeni*, 1-67 (V. Praze, 1933).
- Henson, O. W. *Kans. Univ. Sci. Bull.* **42**, 151-255 (1961).
- Tepaske, E. R. thesis, Oklahoma State Univ. (1964).
- Fleischer, G. *Säugetierk. Mitt.* **21**, 131-239 (1973).
- Fleischer, G. *Evolutionary Principles of the Mammalian Middle Ear* (Springer, Berlin, 1978).
- Henson, O. W. in *Biology of Bats* Vol. 2 (ed. Wimsatt, W. A.), 181-261 (Academic, New York, 1970).
- Novacek, M. J. *Proc. 5th int. Bat Res. Conf.*, 317-330 (Texas Technical Press, Lubbock, 1980).
- Sprague, J. M. *Am. J. Anat.* **72**, 385-472 (1943).
- Griffiths, T. A. *Mammalia* **47**, 377-394 (1983).
- Pye, A. J. *Morph.* **119**, 101-120 (1966).
- Bruns, V. *Verh. dt. zool. Ges.* 141-154 (1979).
- Bruns, V., Fiedler, J. & Kraus, H. J. *Myotis* Vols 21-22 (1983-1984).
- van der Klaauw, C. J. *Bull. Am. Mus. nat. Hist.* **62**, 1-352 (1931).
- Suthers, R. A. & Fattu, J. M. *Am. Zool.* **13**, 1215-1226 (1973).
- von Bekesy, G. *Experiments in Hearing* (McGraw-Hill, New York, 1960).
- Dallos, P. *The Auditory Periphery* (Academic, New York, 1960).
- Harris, D. M. & Dallos, P. *Science* **225**, 741-743 (1984).
- Novacek, M. J. *Proc. 1st int. Congr. Vert. Morphol.* (Fischer, Stuttgart, in the press.)
- Simmons, J. A., Howell, D. J. & Suga, N. *Am. Sci.* **63**, 204-215 (1975).
- Fenton, M. B. *Q. Rev. Biol.* **59**, 33-53 (1984).

Course correction circuitry translates feature detection into behavioural action in locusts

H. Reichert, C. H. F. Rowell & C. Griss

Department of Zoology, University of Basel, Rheinsprung 9, CH-4051 Basel, Switzerland

Like other flying animals, locusts must maintain course stability despite turbulence and motor errors. Accordingly, they must detect course deviations and then correct them. The compound eyes, the ocelli and the cephalic wind hairs all detect different sensory consequences of flight instability¹⁻³, and descending interneurons bring this information to the thorax⁴⁻⁹. One such interneurone of this population (the tritocerebral commissure giant neurone) is known to elicit correctional steering¹⁰. Here we characterize three additional pairs of descending deviation detector neurones and show how their information is translated into altered drive to the flight motoneurons. The central pattern generator for flight¹¹⁻¹³, modulated by proprioceptive feedback¹⁴, gates the signal of the detector neurones in thoracic premotor interneurons, ensuring that the flight motoneurons are affected only during flight. Further, the gating process transforms the phase-independent information of the deviation detectors into a phase-dependent signal modulated at wing-beat frequency, and transfers it to those flight motoneurons active at the time and appropriate to the corrective action required.

We recorded intracellularly, with dye-filled microelectrodes, from the three pairs of descending interneurons and their identified postsynaptic cells in the thoracic ganglia of locusts (*Locusta migratoria* (L. 1758)) during flight motor activity¹², to correlate their morphological and functional characteristics.

Course deviations were simulated by an artificial horizon, provided with a central wind jet (Fig. 1b). In some experiments the ocelli were stimulated independently via light pipes that excluded light from other sources⁸. These procedures established the role of the three interneurons as highly specific absolute deviation detectors, which integrate sensory information from the compound eyes, the ocelli and the wind hairs. Table 1 summarizes the excitatory input characteristics of the three cells.

We have named these interneurons DNI, DNM and DNC, so called because they receive input from the ipsilateral, medial and contralateral ocelli, respectively (Fig. 1a). All have been in part described previously, but because of confusion in the literature (Table 2), we find it necessary to rename them. Input from the compound eye confers on these cells sensitivity to large field movement and to direction of movement. Figure 1c shows an example of the specific nature of compound eye input. A simulated roll deviation presented to the compound eyes results in excitation of the left DNC only if the roll is clockwise and away from the horizontal position. For a roll of this specific type of given amplitude, the response of the cell increases with increasing divergence of the artificial horizon from the horizontal. The cell does not respond to a clockwise roll towards the horizontal, nor does it respond to an anticlockwise roll towards or away from the horizontal. This cell also responds with similar specificity to deviations in the pitch and yaw axes (Table 1). The optimal combination of compound eye stimulation for this cell is what the animal would experience during a diving, banked turn to the side contralateral to the axon.

Excitatory input from the wind hairs or ocelli complements the compound eye input in an aerodynamically compatible and directionally specific manner. For example, a cell which is excited by a visually simulated yaw to the left is also excited by wind directed towards the animal's head from the right which would in nature accompany this yaw. If compound eye input causes the cell to be excited by a simulated clockwise roll, it

Fig. 1 The descending deviation detector neurones DNI, DNM and DNC. **a**, Morphology and nomenclature. The units are so named because they receive input from the ipsilateral, medial and contralateral ocelli, respectively, where ipsi- and contralateral are defined as the side on which the axon of the interneurone runs. This is clearly reflected in the location of the input arborizations in the brain. In the remaining ganglia, DNM is distinguished by its bilateral projection, while DNI and DNC project only unilaterally and are very similar in structure. All three units are paired. We distinguish individual units as left and right with reference to the connectives in which the axon descends to the thorax. Drawings are based on 20-30 successful intracellular fills (either Lucifer yellow³⁵ or hexamine cobalt chloride with subsequent intensification³⁶) of each unit. Scale bar, 800 μ m. **b**, Stimulus apparatus producing simulated course deviations. The animal's head was placed at the centre of a remotely steerable translucent hemisphere, on which was painted the opaque pattern shown, representing an artificial horizon and a structured visual field; the whole was illuminated diffusely with daylight from above and behind. The hemisphere, which was also equipped with a central air tube for simulation of wind on the head in flight, could be rotated around all three orthogonal axes and thus simulate pitch, roll and yaw deviations. **c**, Response of the left DNC to simulated clockwise and anticlockwise rolls away from (upper two records) and towards (lower record) the normal flying position. The sketches show the initial and final positions of the horizon, and the arrows the direction of its rotation. Only clockwise rolls away from the normal position evoke graded responses. Upper traces: intracellular recording, scale bars 50 mV, 200 ms; lower traces: roll monitor, vertical scale bar 75°. **d**, Response of same unit as in **c** to three identical frontal wind stimuli (onset at arrows) during maintained simulated visual rolls to various positions, as shown in the sketches. Calibrations as in **c**. Non-preferred orientations of the visual world inhibit the response to wind. **e**, Response of a DNI (centre trace, intracellular recording, scale bars 5 mV, 200 ms) to a visually simulated roll (lower trace, vertical scale bar 75°) and to simultaneous ON/OFF stimulation of the ipsilateral ocellus by means of a separate light pipe (upper trace, light at 565 nm modulated by 1 log₁₀ unit, upward going = ON, duty cycle 200 ms). Ocellar OFF stimuli evoke only single spikes when presented alone (asterisk). These summate with the response induced by visually signalled roll. Ocellar ON stimuli are functionally incompatible with the visual information, as they imply a roll in the opposite direction; they strongly inhibit the visual response, which otherwise would consist of a long phasic-tonic burst, as in the middle record of **c**.

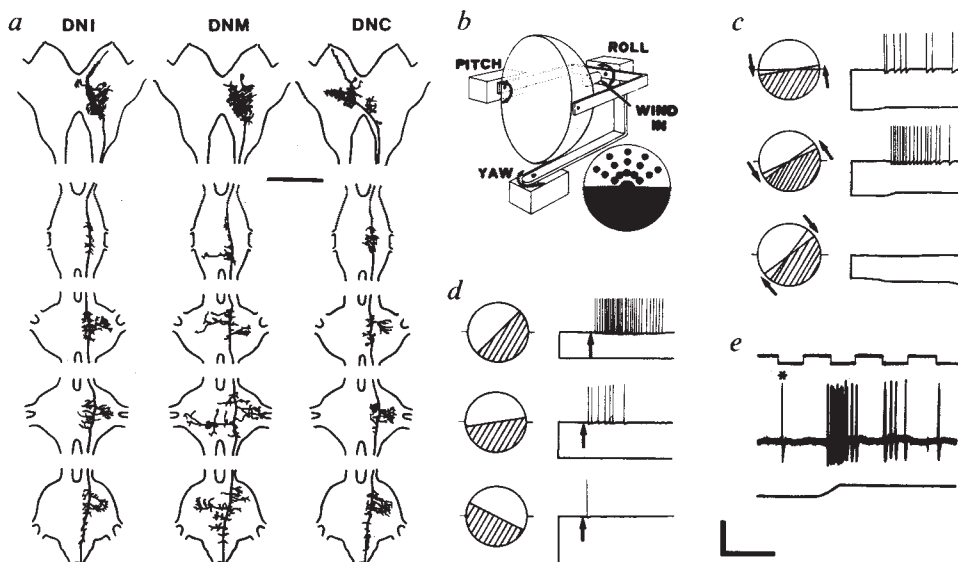


Table 1 Excitatory input characteristics of deviation detector neurones

	DNI	DNM	DNC
Ocellus producing reliable* excitation with OFF stimulation	Ipsilateral	Medial	Contralateral
Most effective visually simulated deviation (that is, via compound eyes)	Roll ipsilateral Pitch down Yaw ipsilateral	Pitch down	Roll contralateral (Pitch down)† Yaw contralateral
Most effective wind direction corresponds to	Yaw ipsilateral	Pitch down	Yaw contralateral

Data were derived from 60–100 experiments on each cell.

* DNI received additionally weak input from the medial ocellus, and DNM from both ipsi- and contralateral ocelli. In a fraction of the population, these weak inputs can evoke spikes.

† Weak response.

will also be excited by selective darkening of the right ocellus, which would in nature be darkened by such a roll¹⁵.

The response properties of a deviation-detecting neurone are not, however, shaped simply by summation of excitation from three sensory sources. Mutual inhibition between the different sensory inputs is also important, and prevents a response if there is a conflict in information content. For example, Fig. 1*d* shows that the position of the visual world can alter the response of the cell to a constant wind stimulus. If the horizon is rolled to a 'visually preferred' position (Fig. 1*d*, top records), indicative of a specific deviation, and held there, the response to a frontal wind stimulus is strong (here 34 action potentials). If the horizon is held in a near-horizontal position (Fig. 1*d*, middle records), the same wind stimulus produces only a modest excitation (7 action potentials). Finally, if the horizon is rolled to a 'visually antipreferred' position, one which does not represent a deviation of interest to this cell, its response to the identical wind stimulus is virtually eliminated (1 action potential). Similarly, Fig. 1*e* shows that the ocellar input can inhibit the response to compound eye input. Darkening of the left ocellus produces an action potential in the left DNI, which responds visually to an anticlockwise roll. However, selective illumination of the same ocellus via a light pipe strongly inhibits the response of the

neurone to an otherwise effective visual roll stimulus. Inhibition of the wind input by ocellar input⁸ and inhibition of the compound eye input by wind input can be demonstrated in similar ways. These inhibitory relations reflect the relative probability of error for a conflict occurring in natural conditions (for example, turbulence could affect the wind hairs without producing a course deviation; the eyes would experience no change in stimulus, and the two inputs to the interneurone would convey conflicting information were it not for the inhibition). Taken together, these interactions ensure that the firing of any of these neurones is a reliable indication of a specific and well-defined deviation from course. Each neurone responds optimally to that specific combination of sensory stimuli which would impinge on the compound eyes, ocelli and wind hairs during a specific deviation. The units are formally multimodal, but we feel that this is not a useful description; the feature detected by these neurones, their 'emergent modality', is course deviation.

The cells central to the transformation of the firing of these deviation detectors into appropriate steering manoeuvres are a population of thoracic interneurones⁹ that have three important characteristics¹⁶ (see Fig. 2*e*): (1) they are postsynaptic to specific descending deviation detectors, (2) they are powerfully presynaptic to specific flight motoneurones and (3) they receive

Fig. 2 Flow of information from the deviation detector neurones (DN) to the flight motoneurones (FMN) is gated in the thoracic interneurones (TIN) by summation with a phasic signal derived from the central pattern generator for flight (CPG). *a*, Presynaptic DNM (dotted) and representative postsynaptic TIN (solid) in mesothoracic ganglion (scale bar, 300 μ m). *b*, Simultaneous intracellular recordings in a non-flying locust (scale bars 2 mV and 50 mV, respectively, 50 ms) from the TIN and DN shown in *a*, and an extracellular recording from the dorsal longitudinal depressor muscle ('emg') which serves as a monitor of flight activity. A depolarizing current of 10 nA ('cm') injected into DN evokes 10 action potentials, which in turn evoke small 1:1 excitatory postsynaptic potentials (e.p.s.ps) in the TIN. (The small high-frequency deflections on the TIN trace are artefacts caused by capacitive coupling between the two electrodes. The input bridges of both recording amplifiers are balanced, but the high current necessary to evoke a brisk burst of action potentials from an axon penetration results in a sizeable stimulus artefact.) *c*, *d*, Same experiment as *b*, but during flight motor output, indicated on the emg trace by periodic depressor muscle potentials. Scale bars 20 mV (TIN), 100 mV (DN) and 100 ms. The TIN receives strong synaptic drive from the CPG, depolarizing in the depressor phase. Note that in *c* and *d* the current pulses come at different points in the wing-beat cycle. During the depressor phase, e.p.s.ps evoked by spikes in the DN increase the number of action potentials in the depolarized TIN by ~50%, which in turn influences the postsynaptic FMN (not shown). Visual stimuli can cause a spike rate in the DN 50% greater than that induced here by current injection, so that this figure underplays the potential effect. During the elevator phase, the same DN activity does not evoke spikes in the TIN, and thus has no effect on the FMN postsynaptic to that cell. *e*, Summary diagram indicating the main relationships in the circuitry discussed here. Thickness of arrows is roughly proportional to the relative strength of the interaction. The TIN population is symbolized as an AND-gate, reflecting its role. CPG drive reaches the FMNs not only via the TINs, but also by other interneurones which do not receive DN input.

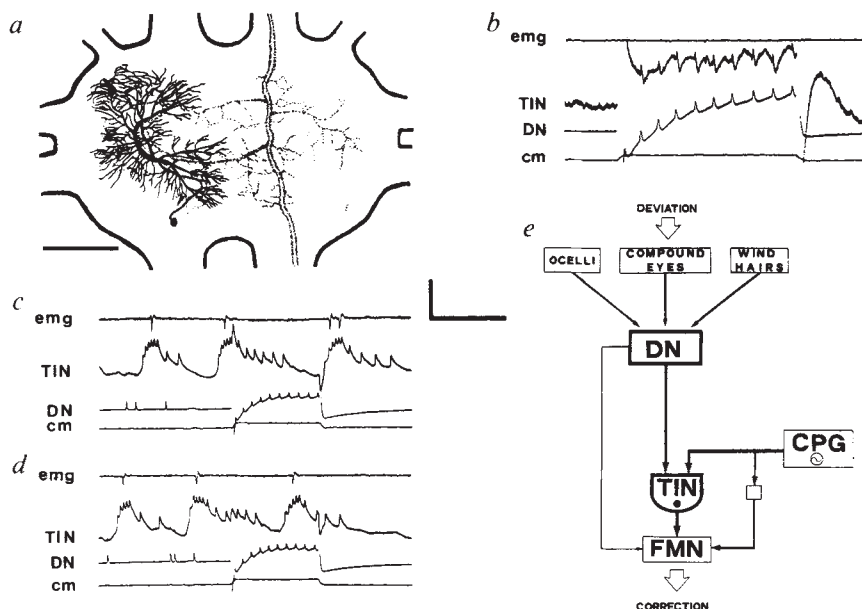


Table 2 Synonymy of the deviation detector neurones

Interneurone			Ref.
DNI	DNC	DNM	This work
O ₃	O ₂	—	28–30
ip 1	cl 1	—	31
O ₂	O ₁	—	32
i.o.	c.o.	—	33
VIP 1	VCP 1	—	6
O ₃	—	O ₃	8
O ₄	—	—	34
Fast	Fast	Fast medial	9
ipsilateral, Fast	contralateral		
ipsilateral (medial)			

Deviation detector neurones have been mentioned under different names by previous authors; all used *Schistocerca* rather than *Locusta*, but we can detect no differences. Their descriptions were based on partial dye fills (for example, thoracic ganglia but not the brain, or vice versa) and either incomplete physiological data or none at all. The only complete structure previously given⁸ is actually derived from two separate units, one filled in the brain and one in the thoracic ganglia. Normally, the oldest nomenclature would be adopted (in this case the O_{1...n} system²⁸). This is no longer practical, as is evident from the table. Abbreviations: O_{2,3,...}, ocellar interneurone 2, 3, ...; ip 1, ipsilateral 3rd-order ocellar neurone; cl 1, contralateral 3rd-order ocellar neurone; i.o., ipsilateral ocellar neurone; c.o., contralateral ocellar neurone; VIP, ventral nerve cord unit, ipsilateral to its soma, protocerebral soma; VCP, ventral nerve cord unit, contralateral to its soma, protocerebral soma.

rhythmically alternating excitatory and inhibitory drive from elements of the flight central pattern generator¹³, at phase angles characteristic for each interneurone. One such thoracic interneurone is shown in Fig. 2a. In a non-flying animal, the firing of the presynaptic DNM evokes only subthreshold synaptic potentials in this thoracic interneurone (Fig. 2b). Postsynaptic flight motoneurons are consequently not affected and no behaviour results. However, in a flying animal the flight central pattern generator is operative and delivers rhythmic drive to the thoracic interneurons. If the deviation detector is now fired (Fig. 2c, d), the synaptic input it evokes in the thoracic interneurone summates with central pattern generator drive, causing the interneurone to fire an increased number of action potentials. This augmentation occurs only when the drive is depolarizing (Fig. 2c, d). Because of the rhythmic nature of the depolarizing and hyperpolarizing potentials evoked by the flight central pattern generator in the thoracic interneurone, the tonic deviation signal is transformed into phasic activity of the premotor interneurons. Sensory information is gated at the level of the thoracic interneurone. The appropriate postsynaptic flight motoneurons will consequently receive additional drive, phase-coupled with the wing-beat cycle, for the duration of the descending deviation signal. This additional phase-coupled motor drive can either alter the phase of firing of one of a pair of homologous bilateral motoneurons, or alter the number of action potentials in the discharge of affected motoneurons, or even recruit an otherwise silent motoneuron. These comprise the mechanisms which have been shown to operate during steering behaviour by intact locusts^{17–20}. All the flight motoneurons and the premotor interneurons are spiking cells with all-or-none threshold characteristics, clearly a prerequisite for this mechanism. We describe elsewhere¹⁶ circuitry in which the CPG-evoked depolarizations in the premotor interneurons summate with inhibitory postsynaptic potentials derived from the deviation detector neurones, rather than with their excitatory postsynaptic potentials, but which has similar effects on the firing pattern of the motoneurons. That summation of inputs to spiking premotor interneurons could lead to gating of the signal to motoneurons has been postulated previously²¹, but not demonstrated.

The deviation detector neurones also make direct connections with the flight motoneurons^{8,9}. We have shown elsewhere¹⁶ that

these connections produce much smaller postsynaptic potentials than those mediated by premotor interneurons. This is understandable in view of the need to avoid flight muscle contraction when the animal is not flying; the direct input to the flight motoneuron must be small enough to avoid firing it, even though the motoneuron receives an almost continuous barrage of sensory excitatory postsynaptic potentials from other sources. In flight, however, these direct postsynaptic potentials are also subject to a similar gating process by the flight central pattern generator, which drives the motoneurons directly¹³ as well as the premotor interneurons.

Thus, at the thoracic level, a two-stage gating process ensures that the intrinsically phase-independent information from the deviation detectors: (1) elicits a response in the flight motoneurons only when the flight central pattern generator is operative, that is, in the behavioural context of flight, (2) is phase coupled, that is, is transformed into a phasic signal modulated at wing-beat frequency and (3) is automatically applied to the appropriate group of flight motoneurons which are active at the time of arrival of the signal, obviating the need for any further decision-making in the brief interval between stimulus and response.

The complete steering behaviour of the animal involves not only the modulation of wing beat treated here, but also movements of head, abdomen and hind legs^{11,22–24}. It remains unclear whether the circuitry described here participates in these elements. However, we have evidence that the information brought to the thoracic ganglia by the other deviation-sensitive interneurons is processed in the same way as described above. Furthermore, the circuit principles documented here seem to be generally applicable and significant. In vertebrates, for example, locomotor control involves descending pathways from projection areas in the brain to segmental motor centres in the spinal cord and thence to spinal motoneurons²⁵. Moreover, a gating process has been postulated and ascribed to the level of the segmental interneurons^{26,27}. This scheme is comparable in detail to the circuitry documented here for the locust (see Fig. 2e). Indeed, we expect that similar principles of neuronal organization will be found in all animals that need to integrate a rhythmic motor output with non-phase-locked exteroceptive information, a need that occurs in most forms of locomotion.

We thank P. Wehrli for technical assistance and W. W. Stewart for a gift of Lucifer yellow. This research was supported by the Schweizerischer Nationalfonds.

Received 23 January; accepted 22 March 1985.

1. Weis-Fogh, T. *Nature* **163**, 873–874 (1949).
2. Goodman, L. J. *J. exp. Biol.* **42**, 385–408 (1965).
3. Taylor, C. P. *J. exp. Biol.* **93**, 1–18 (1981).
4. Camhi, J. M. *J. exp. Biol.* **50**, 349–369 (1969).
5. Kien, J. *Vision Res.* **14**, 1255–1268 (1974).
6. Guy, R. et al. *J. comp. Physiol.* **115**, 337–350 (1977).
7. Bacon, J. & Tyrer, M. *J. comp. Physiol.* **126**, 317–325 (1978).
8. Simmons, P. J. *J. exp. Biol.* **85**, 281–294 (1980).
9. Rowell, C. H. F. & Pearson, K. G. *J. exp. Biol.* **103**, 265–288 (1983).
10. Möhl, B. & Bacon, J. *J. comp. Physiol.* **150**, 453–466 (1983).
11. Wilson, D. M. *J. exp. Biol.* **38**, 471–490 (1961).
12. Robertson, R. M. & Pearson, K. G. *J. comp. Physiol.* **146**, 311–320 (1982).
13. Robertson, R. M. & Pearson, K. G. *J. comp. Neurol.* **215**, 33–50 (1983).
14. Wendler, G. in *Symp. Physiol. Biophys. Insect Flight* Vol. 2 (ed. Nachtigall, W.) 113–125 (Fischer, Stuttgart, 1983).
15. Wilson, M. *J. comp. Physiol.* **124**, 297–316 (1978).
16. Reichert, H. & Rowell, C. H. F. *J. Neurophysiol.* (in the press).
17. Zarnack, W. & Möhl, B. *J. comp. Physiol.* **118**, 215–233 (1977).
18. Taylor, C. P. *J. exp. Biol.* **93**, 19–31 (1981).
19. Wilson, D. M. *J. exp. Biol.* **48**, 631–641 (1968).
20. Waldron, I. *J. exp. Biol.* **47**, 213–228 (1967).
21. Burrows, M. & Horridge, G. A. *Phil. Trans. R. Soc. B* **269**, 49–94 (1974).
22. Goodman, L. J. *J. exp. Biol.* **42**, 385–408 (1964).
23. Camhi, J. M. *J. exp. Biol.* **52**, 519–532 (1970).
24. Gewecke, M. & Phillipson, J. *Physiol. Ent.* **3**, 43–52 (1978).
25. Shepherd, G. M. *Neurobiology*, 395–412 (Oxford University Press, 1983).
26. Forssberg, H., Grillner, S. & Rossignol, S. *Brain Res.* **85**, 103–107 (1975).
27. Grillner, S. *Physiol. Rev.* **55**, 247–304 (1975).
28. Williams, J. L. D. thesis, Univ. Wales (1972).
29. Williams, J. L. D. *J. Zool.* **176**, 67–86 (1975).
30. Goodman, C. S. *Cell Tissue Res.* **175**, 166–183 (1976).
31. Goodman, L. J. in *The Compound Eye and Vision in Insects* (ed. Horridge, G. A.) 515–548 (Clarendon, Oxford, 1975).
32. Patterson, J. A. & Goodman, L. J. *J. comp. Physiol.* **95**, 251–262 (1974).
33. Goodman, L. J. et al. *Cell Tissue Res.* **157**, 467–492 (1975).
34. Simmons, P. J. *J. comp. Physiol.* **147**, 401–414 (1981).
35. Stewart, W. W. *Cell* **14**, 741–759 (1978).
36. Bacon, J. P. & Altman, J. S. *Brain Res.* **138**, 359–363 (1977).

Neuronal growth, atrophy and death in a sexually dimorphic song nucleus in the zebra finch brain

Masakazu Konishi & Eugene Akutagawa

Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

The song control nuclei of the zebra finch brain contain more neurones of larger diameter in the male than in the female¹. This sexual dimorphism is thought to result from differential growth of neurones in the two sexes²⁻⁴. Using neurohistological techniques and radioactive tracers, we have studied the development of several forebrain nuclei involved in the control of song and find that the dimorphism arises from neuronal atrophy and death in the female brain as well as from an increase in cell-body size and afferent terminals from other forebrain nuclei in the male. Although the timing of these events varies from nucleus to nucleus, the sequence is essentially similar in all of them except area X. Here we describe the events in one of these nuclei, the robust nucleus of archistriatum (RA), as an example.

Zebra finches breed well in captivity and provide chicks of known ages all year round. Neurone counts were made on sections stained with cresyl violet and the neuronal connections from HVC (ventral nucleus of hyperstriatum caudal division) to RA were traced by conventional autoradiographical methods. The distinct boundaries of RA are convenient for the measurement of various cytoarchitectonic and cellular properties at all stages of development. For the measurement of RA volume, the boundaries of the nucleus in every third section were drawn with a Zeiss camera lucida onto a sheet of paper and the area within the boundaries was calculated by a computer-aided planimetric method. Because of the error involved in the sampling method, the average between the minimum and maximum estimates of the volume was chosen as the best estimate. The same planimetric method was used for the calculation of soma area from perikaryal outline tracings. For the purpose of this report, the neurone density and number were determined only for a few selected age groups and are not shown in graphical form. Sections having the largest cross-sectional areas of RA were photographed *in toto* and neurones were counted on enlarged prints. The number of neurones counted in each section divided by the volume of the section gave an estimate of the neurone density, which did not differ by more than 10% from that obtained by counting the number of neurones in random sample grids under a microscope. The total number of neurones in RA was estimated by multiplying the neurone density by the RA volume.

The cytoarchitecture of RA reveals few differences between the sexes soon after hatching but later begins to diverge towards growth in the male and atrophy in the female (Figs 1, 2). The growth in RA volume in the male is mainly the result of changes in the size and density of neuronal cell bodies; the somata become larger and less densely packed. The estimate of neurone number at day 15 ranges from 9,100 to 14,900 (median = 11,300) and from 9,200 to 14,900 (median = 9,600), respectively, for the male and the female. These numbers should be compared with those of adult birds which range from 9,300 to 14,300 (median = 13,600) for the male and from 2,400 to 3,700 (median = 2,900) for the female. Thus, the female RA loses about two-thirds of its neurones by the adult stage, whereas the male RA retains essentially the same range of neurone numbers. The sharp decline of RA volume in the female after day 30 results from both a decrease in the neurone number and size on one hand and an increase in the neurone density on the other; the cells are fewer, smaller and more densely packed. These small cells were identified as neurones by retrograde labelling from the target areas of RA neurones³.

During their development RA neurones are mainly innervated by axons from HVC and to a lesser extent by axons from MAN

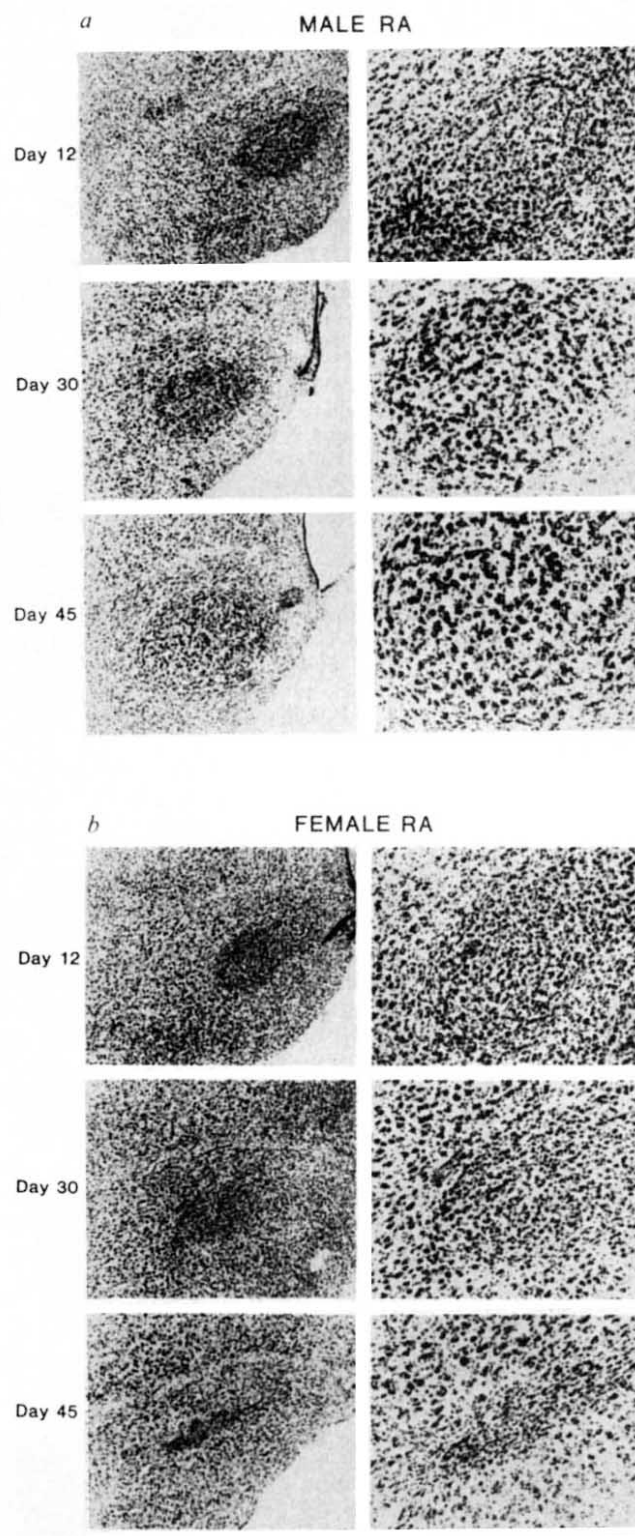
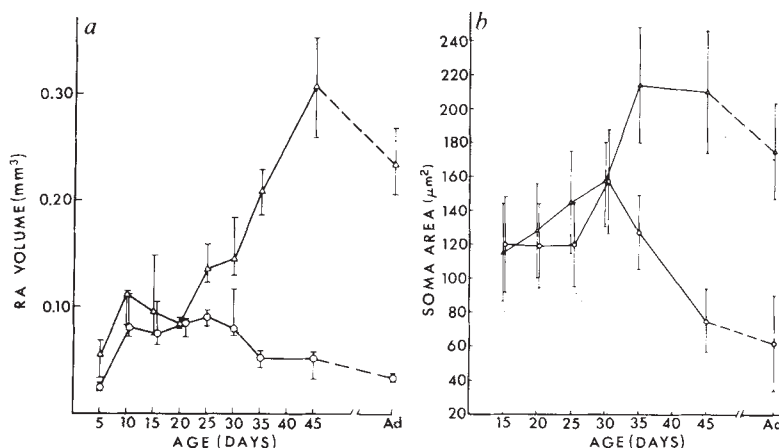


Fig. 1 Contrasting developmental stages of RA in the two sexes. At day 12 after hatching, the male (a) and female (b) RAs are similar in size; at day 30, the cell bodies of neurones are smaller in the female RA than in the male RA. By day 45 the female RA has shrunk, whereas the male RA has grown both in volume and soma size. For histological examination of brains, birds were anaesthetized with Equithesin, perfused with saline and then 10% formalin. After overnight immersion in 20% sucrose solution, brains were cut in 30 μ m thickness with a freezing microtome, and these sections were stained with cresyl violet. Each photomicrograph shows the largest cross-sectional area of RA for a given age in a parasagittal section. Low- and high-magnification pictures of the same section are shown side by side. Scale bars, 100 μ m.

Fig. 2 Quantitative comparison of architectonic and neuronal changes (male, Δ ; female, $\circ$). Dashed lines indicate unsampled age intervals. *a*, Changes in RA volume. For each age and sex group, the median and range of the RA volume of five birds are plotted as a function of age. From days 5 to 20 after hatching, RA grows to about an equal volume in both sexes. Thereafter, the male RA continues to grow with a steep increase in growth rate at about day 30, although there is a temporary dip in the growth curve at about day 20. The volume of the female RA, on the other hand, remains almost constant from day 10 to day 30 when a sharp decline begins. *b*, Changes in neurone size. Each of five birds used for each age and sex group provided a sample of 50 neurones for the computation of a mean and standard deviation. The graph plots the means of the five means and standard deviations as a function of age. In the male, the mean cross-sectional area of RA neurones increases steadily, reaching a plateau at about day 35. In the female, the mean soma area remains more or less constant until about day 30 when it begins to decrease rapidly.



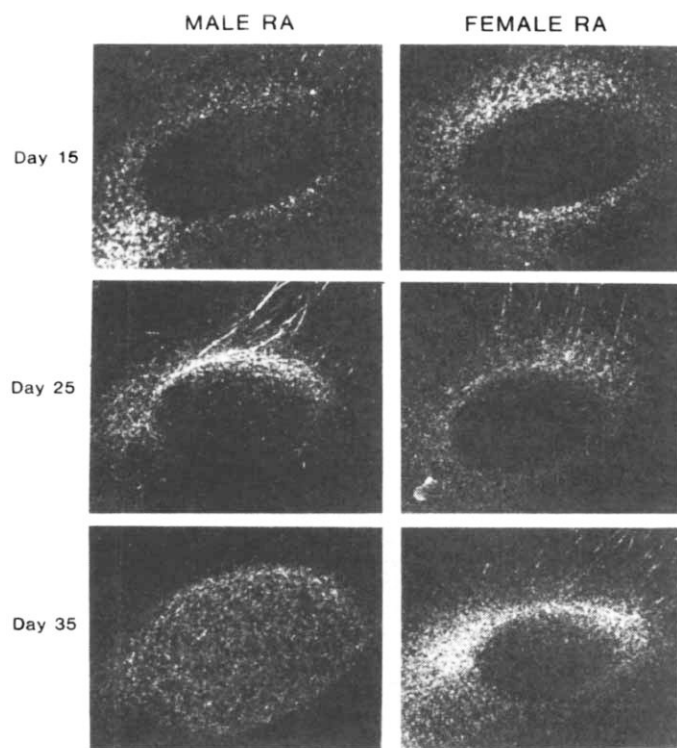
(magnocellular nucleus of anterior neostriatum). The time course and pattern of innervation also show differences between the sexes. An injection of tritiated proline in HVC and its immediate vicinity in 15–25-day-old chicks of either sex results in an accumulation of radioactivity (grains) around RA (Fig. 3), while control injections outside HVC and its adjoining areas do not produce this pattern of grain distribution. Furthermore, when radioactivity accumulates within the boundaries of RA as in males older than 35 days, the ring of grain density around the nucleus disappears (Fig. 3). Whatever the identity of the ring may be, labelling within the boundaries of RA indicates innervation of its neurones. Thus, the male RA appears to remain uninnervated until about day 30 when it suddenly becomes innervated. This stage corresponds to the beginning of the steep increase in RA volume. The soma density begins to decrease rapidly at about this age, presumably because the accommodation of the incoming afferent fibres results in an increase in the distance between RA neurones. The female RA appears to remain largely uninnervated, although sparse connections between HVC and RA may present in the adult³.

The above findings indicate that cell atrophy and death constitute a normal process of sexual differentiation in the brain.

In theory, differences in neurone number and size in sexually dimorphic areas of the brain may result from differential proliferation, growth, atrophy, death or migration^{5,6}. However, we did not consider cell atrophy and death in our earlier work on the zebra finch. Consequently, we did not consider that early oestrogen may masculinize the female RA by preventing cell death²⁻⁴. As to cell migration, it is very unlikely that large neurones emigrate from the female RA and small neurones migrate into it. Neurogenesis is reported to occur in the adult zebra finch brain⁷. Because neurogenesis in the adult is generally not restricted either to the song-control system or to one sex, it is perhaps not a major factor in the sexual differentiation of the song-control system. The presence in the male and the absence in the female of neuronal connections between HVC and RA suggest a causal relationship between innervation and neuronal growth: the afferent innervation of RA may be essential for the growth of its neurones; alternately, the growth of RA neurones may induce innervation by HVC fibres. How oestrogen mediates one of these processes is not known.

The origin of brain sex differences by neuronal growth in the male and cell death in the female is not unique to the zebra finch, for all well-documented cases of brain sex differences appear to involve cell death in the female. Recent evidence indicates that differences in neurone populations between the sexes in the preoptic area and the bulbocavernosus nucleus of mammals involve growth of neurones in the male and cell death in the female⁸⁻¹¹. Cell death is a common phenomenon in the development of the nervous system; however, the role of cell death in development varies from system to system¹². For sexually dimorphic systems the significance of cell death is clear; neurones disappear in the female system because females do not need to exhibit the behaviour controlled by the relevant neurones. Why are the neurones that are destined to die produced in the first place? It could be that a genetic programme uses the same set of genes for a homologous neural circuit in both sexes and an epigenetic process determines the fate of the circuit in the two sexes.

Fig. 3 Sex differences in innervation of RA. A glass pipette, 15 μ m in tip diameter, was used for the iontophoretic injection of tritiated proline into HVC. The survival time after the tracer injection was 3 days. Brain sections for autoradiography were exposed for 30 days at 4 °C. Dark-field autoradiographs of RA show the distribution of silver grains for three different ages. These patterns of grain distribution appear when tritiated proline is injected in HVC and its immediate vicinity. The unlabelled oval area is RA. At day 15 both male and female RA fail to accumulate radioactivity no matter how large the injection area may be. At day 25 the male RA shows some labelling within its boundaries and is completely labelled by day 35. Note the absence of labelling outside the boundaries of RA. In contrast, the female RA remains mostly unlabelled throughout its development including the adult stage. These changes in labelling indicate when and to what extent RA becomes differently innervated in the two sexes. Scale bar, 100 μ m.



We thank Drs C. E. Carr and P. Patterson for reading the manuscript and Chris Yo for assistance. This work was supported in part by a grant from the Whitehall Foundation.

Received 26 November 1984; accepted 11 February 1985.

1. Nottebohm, F. & Arnold, A. P. *Science* **194**, 211–213 (1976).
2. Gurney, M. E. & Konishi, M. *Science* **208**, 1380–1383 (1980).
3. Gurney, M. E. *J. Neurosci.* **1**, 658–673 (1981).
4. Konishi, M. & Gurney, M. E. *Trends Neurosci.* **5**, 20–23 (1982).
5. Goy, R. W. & McEwen, B. S. *Sexual Differentiation of the Brain* (Massachusetts Institute of Technology Press, 1980).
6. Arnold, A. P. & Gorski, R. A. *Rev. Neurosci.* **7**, 413–742 (1984).
7. Nottebohm, F. *Condor* **86**, 227–236 (1984).
8. Breedlove, S. M. *Soc. Neurosci. Abstr.* **10**, 927 (1984).
9. Sengelaub, D. R. & Arnold, A. P. *Soc. Neurosci. Abstr.* **10**, 453 (1984).
10. Nordeen, E. J., Nordeen, K. W., Sengelaub, D. R. & Arnold, A. P. *Soc. Neurosci. Abstr.* **10**, 453 (1984).
11. Ulibarri, C. & Yahr, P. *Soc. Neurosci. Abstr.* **10**, 454 (1984).
12. Cowan, W. M., Fawcett, J. W., O'Leary, D. D. M. & Stanfield, B. B. *Science* **225**, 1258–1265 (1984).

Changes of free calcium levels with stages of the cell division cycle

Martin Poenie*, Janet Alderton†, Roger Y. Tsien* & Richard A. Steinhardt†

* Departments of Physiology-Anatomy and † Zoology, University of California, Berkeley, California 94720, USA

Although the regulation of events in the cell division cycle by calcium or other cations has been the subject of much interest and speculation^{1–10}, experimental studies have been hampered by the difficulty of measuring submicromolar intracellular free calcium concentrations ($[Ca^{2+}]_i$) over an entire cell cycle. We now describe experiments using a new fluorescent calcium chelator, fura-2 (see Fig. 1c for structure), for continuous measurement of $[Ca^{2+}]_i$ from fertilization through the first cleavage of individual eggs of the sea urchin *Lytechinus pictus*. We also show for comparison the results of parthenogenetic activation by ammonia. In addition to the known transient rise of $[Ca^{2+}]_i$ at fertilization, further peaks are now revealed during pronuclear migration, nuclear envelope breakdown, the metaphase/anaphase transition and cleavage. Parthenogenetic activation by ammonia also elicits a sustained rise starting at nuclear envelope breakdown.

Fura-2 is a fluorescent tetracarboxylate chelator that exhibits a spectral shift in excitation maxima on binding calcium¹¹. Increasing free Ca^{2+} increases the fluorescence excited by 350 nm light and decreases the fluorescence from 385 nm excitation. The ratio of fluorescence excited by 350 nm to that excited by 385 nm is a measure of $[Ca^{2+}]_i$, in which the uncertainties in dye concentration, cell thickness and absolute optical efficiency of the instrument are cancelled. We measured 350/385-nm ratios from single eggs, injected with fura-2, on a fluorescence microscope that had rapidly alternating excitation wavelengths¹². Ratios were calibrated in terms of $[Ca^{2+}]_i$ by comparison with dye samples of known $[Ca^{2+}]_i$. Figure 1a shows the fluorescence signal during fertilization, recorded simultaneously at 350 nm and 385 nm excitation wavelengths. Shortly after addition of sperm, the fluorescence from 385 nm excitation fell dramatically whereas the fluorescence from 350 nm increased. Changes in the 350/385-nm fluorescence ratio are shown in Fig. 1b, while Fig. 1c shows the actual $[Ca^{2+}]_i$ after dye calibration. From these records we calculated that resting $[Ca^{2+}]_i$ averages 144 ± 10.3 nM (mean $\pm$ s.e.m., $n = 11$), with extremes between 70 and 250 nM (data not shown). The fertilization transient takes ~ 12 s from beginning to peak and rises to an average of 1.95 ± 0.16 μ M ($n = 11$). Levels as high as 3.5 μ M were observed, although such eggs were usually polyspermic. These values are close to those estimated previously from fertilization experiments with aequorin-loaded eggs¹³. The fertilization transient lasts ~ 5 min, during which time the sperm is incorporated into the egg and the sperm nucleus is transformed into the male pronucleus^{14,15}.

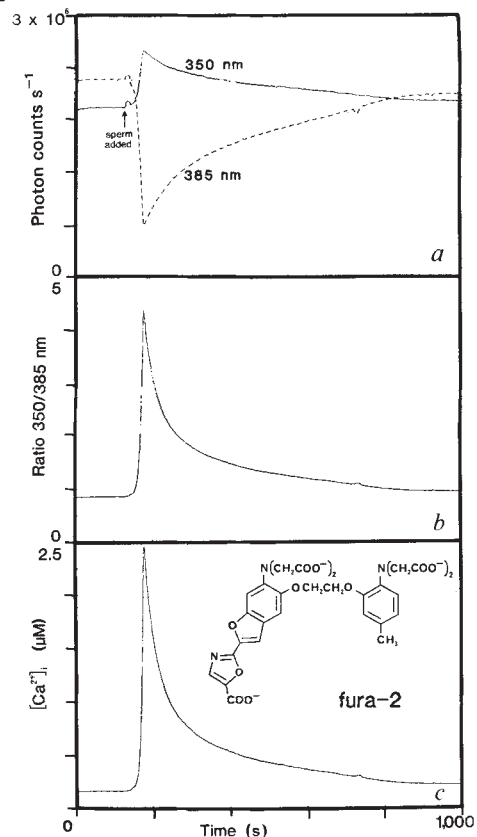


Fig. 1 *a*, Fluorescence signal at 350 nm (solid line) and 385 nm (dotted line) excitation during fertilization, ending just before the onset of a second (pronuclear migration) peak. The addition of sperm (at the arrow) caused a small artefact which can be seen to cancel out when ratios are taken (see *b*). Shortly after the addition of sperm, the fluorescence at 385 nm decreases dramatically (broken line) whereas the signal increases at 350 nm (solid line). *b*, Fluorescence ratio (R), obtained by dividing the fluorescence values at 350 nm by those at 385 nm. This ratio changes in the same direction as $[Ca^{2+}]_i$ and can be used to calculate $[Ca^{2+}]_i$ from the equation¹¹ $K(R - R_0)/(R_s - R) = [Ca]$, where R_0 is the ratio at 0 calcium and R_s is the ratio at saturating calcium. K represents $K_d(F_0/F_s)$, where K_d is the effective dissociation constant for fura-2 in the appropriate conditions, F_0 is the fluorescence at 385 nm in 0 Ca and F_s is the fluorescence at 385 nm in saturating calcium solutions. R_0 and R_s were obtained from droplets of fura-2 reference standards containing 50 μ M fura-2, 155 mM KCl, 25 mM NaCl, 100 mM MOPS neutralized with KOH to pH 7.02 and either 10 mM K_2H_2 -EGTA or 1 mM $CaCl_2$. The effective dissociation constant for fura-2 was determined to be 7.74×10^{-7} M by cuvette titration of 5 μ M fura-2 in the above KCl/NaCl/MOPS medium with 10 mM EGTA buffers. The necessary dissociation constants for Ca-EGTA buffers were measured at 18°C in 225 mM KCl and 25 mM NaCl (ref. 11). *c*, Actual $[Ca^{2+}]_i$ values derived from the ratio in *b* after applying the calculation shown above. The inset shows the structure of fura-2.

Methods. *Lytechinus pictus* eggs were microinjected with a solution of 40 mM fura-2, 5 mM $CaCl_2$ and 0.66 M potassium gluconate (the injection solution was prepared by hydrolysing the fura-2 pentaethyl ester in 1 M KOH and neutralizing the hydrolysate with D-gluconic acid lactone). Typical dye loadings were 50–200 μ M based on comparison with 100- μ M droplets of dye standards. Excitation light was provided by two Spex Industries monochromators preset at 350 nm and 385 nm and selected in rapid alternation (15 Hz) by a rotating chopper mirror. These wavelengths were displaced from the true excitation peaks to avoid the transmission cutoff of glass optics at lower wavelengths. Emitted light was collected through a 500–530-nm interference filter and photon-counted in synchrony with the chopper so that the signals from the two excitation wavelengths were stored in separate memories of a Spex Datamate microcomputer. Monochromator slits of 0.25 mm were used, giving a 0.9-nm bandpass, and an attenuator screen further reduced the excitation intensity fivefold. Dye bleaching was usually undetectable, yet the photometer count rate was $1\text{--}2 \times 10^6$ photon counts per s; uninjected eggs gave $0.5\text{--}1.5 \times 10^4$ counts per s. All experiments were at 18°C.

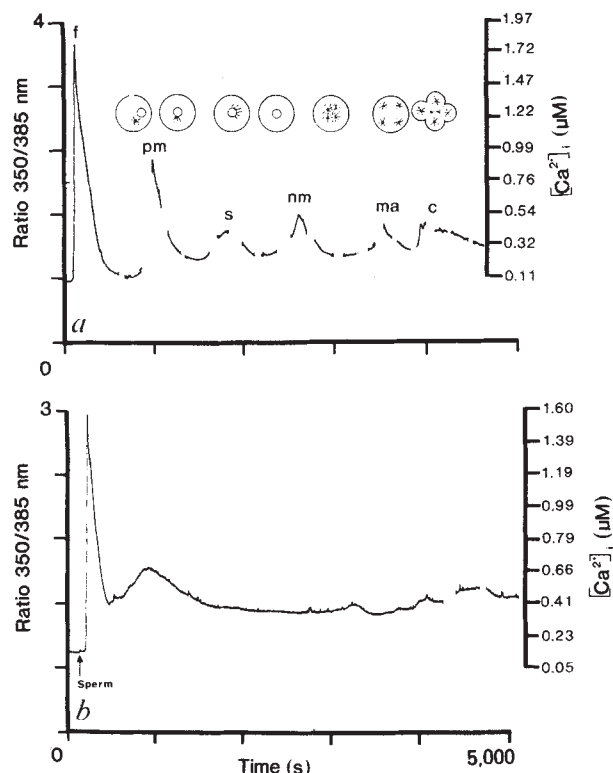


Fig. 2 Two complete records from fertilization to the first cell division. **a** Shows clearly the pattern of calcium fluctuations during the cell cycle seen in our records. Peaks in this record are especially prominent, as the post-fertilization baseline $[Ca^{2+}]_i$ is unusually low. However, the peak calcium levels still attain normal values, similar to those of the corresponding periods of **b**. The breaks in the trace were periods in which the trace was interrupted in order to take transmitted light photographs (not shown) to record any corresponding morphological events: f, fertilization peak; pm, pronuclear migration peak; s, streak stage peak; nm, nuclear membrane breakdown peak; ma, metaphase/anaphase peak; c, cleavage peak. This example happened to be polyspermic, and the first division resulted in four cells. **b** Is a typical trace where the post-fertilization baseline is higher. This egg underwent normal cleavage between the first and second breaks at the end of the record.

The low illumination intensities required with fura-2 and its apparent low toxicity allowed continuous measurement of calcium throughout the cell cycle. Embryos routinely cleaved at least three times (see Fig. 1 legend for details). Figure 2 shows two examples of continuous records obtained from fertilization to first cleavage. Cyclical changes in $[Ca^{2+}]_i$ coincide with specific events of embryo development and the cell division cycle. Figure 2a shows particularly well-defined $[Ca^{2+}]_i$ peaks corresponding to different stages of the cell cycle. In the most typical records (see Fig. 2b), peaks reach values similar to those in Fig. 2a but appear less striking because they rise from a permanently elevated plateau level of $[Ca^{2+}]_i$. Both records reveal that at 10–12 min after fertilization, a second, surprisingly large, increase in $[Ca^{2+}]_i$ occurs. Photographs taken during the breaks in Fig. 2a, as well as many observations during other experiments, indicate that this peak occurs during pronuclear migration. The average peak $[Ca^{2+}]_i$, 665 ± 51 nM ($n = 8$), is well above the mean post-fertilization baseline of 277 ± 28 nM ($n = 8$). EGTA-induced reduction of the external calcium concentration to 10^{-7} M did not affect this peak or subsequent peaks, suggesting that the increases in calcium result from release from internal stores (data not shown).

In 10–20% of the embryos (including that of Fig. 2a), we observed another calcium peak associated with the streak stage (the period of pole separation and DNA synthesis)¹⁶. This transient was seen in both polyspermic and monospermic embryos.

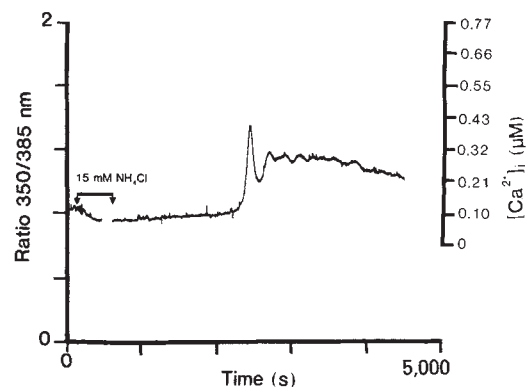


Fig. 3 A typical example of an egg activated by a 10-min pulse (arrows) of sea water containing 15 mM ammonium chloride, pH 8.0. From previous measurements²¹, intracellular pH would be expected to rise from 6.9 towards 7.4–7.6 during ammonia exposure and to remain permanently above 7.3 even after ammonia withdrawal. Calcium levels can be seen to fall slightly with the addition of ammonia, but to rise sharply at 40 min after the addition. These results are typical of three similar experiments. Further experiments (not shown) give similar results even when ammonia is left in continuously or when external calcium is reduced to 10^{-7} M.

Another elevation of $[Ca^{2+}]_i$ occurred 50–60 min after fertilization, with an average peak value of 404 ± 20 nM ($n = 8$). Photographs indicate that the increase in $[Ca^{2+}]_i$ occurs close to or during nuclear envelope breakdown. This peak in free calcium may explain earlier findings of Bennet and Mazia¹⁷ in which unfertilized eggs that had been fused with fertilized embryos underwent cortical granule breakdown in the unfertilized half when the fertilized half began mitosis. The trigger for Ca^{2+} at nuclear envelope breakdown may have triggered the release of Ca^{2+} stores in the unfertilized half.

A fourth peak was observed ~70 min after fertilization, with a mean value of 410 ± 10 nM ($n = 6$). Photographs and visual checks suggest that this peak occurs during the metaphase/anaphase transition, although precise timing is difficult because we cannot resolve the actual chromosome movements. However, in Fig. 2a, the spindle was newly formed before the peak, and spindle elongation, corresponding to late anaphase, did not occur until after the peak, thus indicating that the metaphase/anaphase transition occurred during the period of the peak. Although other investigators have suggested that calcium may be involved in the metaphase/anaphase transition, until now there has been no evidence that an increase in $[Ca^{2+}]_i$ actually takes place. Izant has reported that injections of calcium buffered at 1 μ M or 50 nM free calcium, respectively, shorten or lengthen the duration of metaphase⁴. Wolniak *et al.* reported a decrease in chlortetracycline staining during metaphase, suggesting that a release of calcium stores occurs at the onset of anaphase⁷.

Finally, we observed an increase in $[Ca^{2+}]_i$ associated with cleavage. We found that calcium rises to a mean value of 412 ± 34 nM ($n = 7$); this increase slightly precedes and occurs in conjunction with the developing cleavage furrow. The notion that calcium may regulate the activity of the contractile ring has been suggested by several investigators using injection of calcium buffers or ionophore A23187 (refs 1, 2, 6), although attempts to measure a change with aequorin have been equivocal^{1,18}. On the other hand, $[Ca^{2+}]_i$ measurement in *Xenopus* embryos using ion-selective microelectrodes consistently failed to show evidence of cell-cycle-related fluctuations in calcium¹⁹; this was perhaps a result of species differences or localization of calcium rises that always missed the electrode tip.

The general elevation of baseline $[Ca^{2+}]_i$ and upward fluctuations during the cell cycle are not the only ionic events that follow fertilization of sea urchin eggs. A rise in internal pH, normally dependent on external Na^+ , is known to be necessary for development^{20,21}. The addition of ammonia to sea water, which artificially elevates the internal pH, induces many of the

events of the cell cycle, such as DNA synthesis, chromosome condensation and nuclear envelope breakdown²². We hypothesized that if calcium is pertinent to these cell-cycle events, ammonia might affect $[Ca^{2+}]_i$; however, at first sight this seems unlikely as ammonia does not cause a cortical reaction or any obvious change in NADP/NAD ratios, both of which are calcium-dependent consequences of fertilization^{23,24}. We measured $[Ca^{2+}]_i$ in ammonia-treated eggs (Fig. 3) and found that during the first 40 min after addition of ammonia, $[Ca^{2+}]_i$ either decreased slightly or did not change. A decrease in $[Ca^{2+}]_i$ would be interesting in light of the finding that the activity of the calcium-sequestering system is enhanced by elevated pH²⁵. However, at ~40–55 min, $[Ca^{2+}]_i$ rose to levels characteristic of fertilized eggs entering mitosis (350–400 nM) and remained elevated thereafter. Inspection revealed that the nuclear envelope broke down at 60–70 min, shortly after the increase in $[Ca^{2+}]_i$. Similar results were obtained even if external calcium was lowered to 10^{-7} M, indicating that ammonia can eventually reset the intracellular calcium levels even in the absence of external calcium. This finding requires a revision of the conclusions reached by Zucker *et al.* on the action of ammonia on the nuclear envelope²⁶. Using aequorin, they were unable to detect these low calcium concentrations during the ammonia treatment and so concluded that ammonia exerted its effect through alkalization alone. Although ammonia does raise $[Ca^{2+}]_i$ after 40–55 min, earlier events, such as the acceleration of protein synthesis which starts 25 min after ammonia addition, must be pH responses not mediated by calcium²⁷.

The observed correlation between mitotic events and changes in $[Ca^{2+}]_i$ does not prove that these changes are required for

the observed events. The causal relationship between ionic and mitotic events will be clarified when a means becomes available for controlling $[Ca^{2+}]_i$ at given times.

This work was funded by NSF grants PCM81-04294 and PCM84-02089 to R.A.S. and NIH grants GM31004 and EY04372, and the Searle Scholars Program (83-K-111) to R.Y.T. We thank Drs Zac Cande and Daniel Mazia for valuable discussions.

Received 30 October 1984; accepted 25 February 1985.

1. Baker, P. F. & Warner, A. E. *J. Cell Biol.* **53**, 579–581 (1972).
2. Berridge, M. J. *Symp. Soc. exp. Biol.* **30**, 219–231 (1976).
3. Gingell, D. J. *J. Embryol. exp. Morph.* **23**, 583–609 (1970).
4. Izant, J. G. *Chromosoma* **88**, 1–10 (1983).
5. Means, A. R. & Dedman, J. R. *Nature* **285**, 73–77 (1980).
6. Schroeder, T. E. & Strickland, D. L. *Expl. Cell Res.* **83**, 139–142 (1974).
7. Wolniak, S. M., Hepler, P. K. & Jackson, W. T. *J. Cell Biol.* **96**, 598–605 (1983).
8. De Laat, S. W., Buwalda, R. J. A. & Habets, A. M. M. C. *Expl. Cell Res.* **89**, 1–14 (1974).
9. Lee, S. L. & Steinhardt, R. A. *J. Cell Biol.* **91**, 414–419 (1981).
10. Morisawa, M. & Steinhardt, R. A. *Expl. Cell Res.* **140**, 341–351 (1982).
11. Grynkiewicz, G., Poenie, M. & Tsien, R. Y. *J. biol. Chem.* **260**, 3440–3450 (1985).
12. Tsien, R. Y., Rink, T. J. & Poenie, M. *Cell Calcium* (in the press).
13. Steinhardt, R. A., Zucker, R. & Schatten, G. *Dev. Biol.* **58**, 185–196 (1977).
14. Long, F. J. & Anderson, E. *J. Cell Biol.* **39**, 339–368 (1981).
15. Bestor, T. H. & Schatten, G. *Dev. Biol.* **88**, 80–91 (1981).
16. Harvey, E. B. *The American Arbacia and other Sea Urchin Eggs* (Princeton University Press, 1956).
17. Bennet, J. & Mazia, D. *Expl. Cell Res.* **134**, 494–498 (1981).
18. Gilkey, J. C., Jaffe, L. F., Ridgway, E. B. & Reynolds, G. T. *J. Cell Biol.* **76**, 448–466 (1978).
19. Rink, T. J., Tsien, R. Y. & Warner, A. E. *Nature* **283**, 658–660 (1980).
20. Johnson, J. D., Epel, D. & Paul, M. *Nature* **272**, 661–664 (1976).
21. Shen, S. S. & Steinhardt, R. A. *Nature* **272**, 253–254 (1978).
22. Epel, D., Steinhardt, R. A., Humphreys, T. & Mazia, D. *Dev. Biol.* **40**, 245–255 (1974).
23. Epel, D., Patton, C., Wallace, R. W. & Cheung, G. *Cell* **23**, 543–549 (1981).
24. Whitaker, M. J. & Steinhardt, R. A. *Cell* **25**, 95–103 (1981).
25. Suprynowicz, F. A., Poenie, M. & Mazia, D. *J. Cell Biol.* **97**, 30a (1983).
26. Zucker, R. S., Steinhardt, R. A. & Winkler, M. M. *Dev. Biol.* **65**, 285–295 (1978).
27. Winkler, M. M., Steinhardt, R. A., Grainger, J. L. & Minning, L. *Nature* **287**, 558–560 (1980).

Transformation of mononuclear phagocytes *in vivo* and malignant histiocytosis caused by a novel murine spleen focus-forming virus

T. Franz*, J. Löhler*, A. Fusco†, I. Pragnell‡, P. Nobis§, R. Padua|| & W. Ostertag*

* Heinrich-Pette-Institut für Experimentelle Virologie und Immunologie an der Universität Hamburg, Martinistrasse 52, 2000 Hamburg 20, FRG

† Dipartimento di Biologia e Patologia Cellulare e Molecolare, Facoltà di Medicina e Chirurgia dell'Università di Napoli, via S. Pansini 5, 80131 Naples, Italy

‡ The Beatson Institute for Cancer Research, Bearsden, Glasgow G61 1BD, UK

§ Institut für Physiologische Chemie, Universität Hamburg, 2000 Hamburg 13, FRG

|| Marie Curie Memorial Foundation, Research Institute, The Chart, Oxted RH8 0TL, Surrey, UK

The study of retrovirus-induced leukaemias in mice is a powerful tool for the elucidation of the normal regulation of the haematopoietic system. The acute murine spleen focus-forming viruses (SFFV) can be classified according to the haematopoietic lineage on which they exert their effects in the adult mouse. Here we report a new SFFV isolate, the AF-1 virus, with the novel ability to transform cells of the mononuclear phagocyte lineage. The virus was isolated from sarcomas that were induced on passage of a cloned Friend helper virus (F-MuLV, 643/22F)¹ in newborn BALB/c mice. We have cloned the transforming defective subunit of the AF-1 viral complex in NRK cells and isolated several subclones. Analysis of the proviral genome in two non-producer cell clones reveals that AF-1 virus contains Harvey *v-ras*-specific sequences (Fig. 1). Thus, AF-1 virus is closely related to Harvey murine sarcoma virus (Ha-MSV)², and is, at present, the only tool by which permanent cell lines can be obtained from mononuclear phagocytes in the mouse.

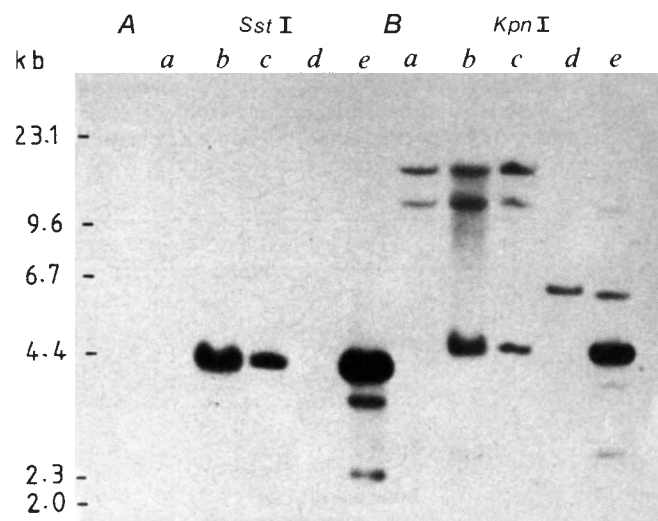


Fig. 1 Southern blots showing Ha-ras sequences. *SstI*(A) and *KpnI*(B)-digested high-relative molecular mass DNA was electrophoresed in horizontal 0.8% agarose gels and transferred to nitrocellulose filters¹⁷. The filters were hybridized in stringent conditions¹⁸ to 10^6 c.p.m. ml⁻¹ of ³²P-labelled clone BS9 of v-Ha-ras (specific activity 2×10^8 c.p.m. μ g⁻¹)¹⁹. Lanes a, NRK; b, AF124T; c, AF124Tc13; d, 3T3; e, HT3. *HindIII*-digested c1857 DNA served as relative molecular mass markers: kb, kilobases. AF124T is an NRK non-producer cell clone of AF-1 virus (see text). AF124Tc13 is a subclone of AF124T. HT3 is a Ha-MSV non-producer cell clone²⁰.

Intravenous injection of AF-1 virus rapidly caused splenomegaly and anaemia in adult DDD (*Fv-2*^s) (Table 1) and B6.S mice (data not shown). DDD mice infected with AF-1 virus died within the first 25 days of the disease. C57BL/6 mice of *Fv-2*^s or *Fv-2*^r genotype were more resistant than DDD mice. C57BL/6 *Fv-2*^r mice survived for up to 50 days after infection. The relative resistance of C57BL/6 mice was also shown using spleen focus formation as the parameter of viral host range (T.F. *et al.*, unpublished observations).

The histological examination of AF-1 virus-infected DDD (*Fv-2^s*) mice revealed that the rapidly progressing spleen enlargement results from the growth of a tumour of the monocyte-macrophage/histiocyte lineage (Fig. 2). Transformed histiocytic cells were first detected in the marginal zone of the spleen 6 days after infection and then invaded the adjacent red and white pulp. At later stages these tumour cells were also found in bone marrow, lymph nodes, Peyer's patches and many parenchymal organs, but not the thymus or the brain. During the course of the disease, the cellularity of the bone marrow decreased 10-fold and at later stages the tissue recovered from the femoral cavity consisted mainly of histiocytic tumour cells. The disseminated and invasive growth of the malignant histiocytic neoplasm and the pronounced pancytopenia in AF-1 virus-infected mice were strikingly reminiscent of human malignant histiocytosis (also known as malignant medullary reticulosis), which is a systemic and fatal histiocytic tumour of both children and adults³⁻⁶.

We also studied the effects of AF-1 virus on haematopoietic colony-forming cells (Table 1): 18 days after infection of DDD mice with AF-1 virus, no early erythroid precursor cells (burst-forming units—erythroid, BFU-E) could be detected in the bone marrow. Late erythroid precursor cells (colony-forming units—erythroid, CFU-E) were markedly decreased in the bone marrow and the spleen. The disappearance of erythroid colony-forming cells in the haematopoietic tissues of AF-1 virus-infected mice correlated well with the progressive decrease of erythroid precursor cells in the histological specimens.

The myeloid/monocytic precursor cells (CFU-C), however, were increased in the spleen of AF-1 virus-infected mice and most of them proliferated in the absence of stimulating factors. At other time points after infection with AF-1 virus, DDD (*Fv-2^s*) mice also showed factor-independent CFU-C in the bone marrow. In our culture system this factor-independent CFU-C formation from spleen cells of AF-1 virus-infected mice was linear, with plating densities ranging from 10^4 to 10^5 cells per dish, indicating that the factor independence is not related to the presence of factor-producing accessory cells. We obtained

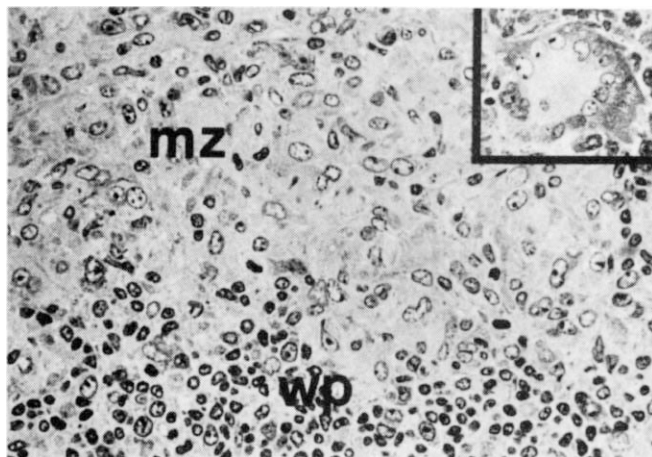


Fig. 2 Spleen of DDD mouse 10 days after infection with AF-1 virus. The normal cellular composition of the marginal zone (mz) of the white pulp (wp) is replaced by cohesively growing histiocyte-like tumour cells spreading into the red and white pulp. The tumour cells are characterized by irregular rounded, elongated, reniform or lobulated nuclei of variable size with finely dispersed chromatin and one or several prominent nucleoli. They possess abundant pale eosinophilic cytoplasm with irregular blurred cell borders. In poorly differentiated tumour cells, strong basophilic staining of the cytoplasm is seen. Bi- and multinucleated cells (inset) and mitotic figures are present. Glycol-methacrylate embedding, Giemsa stain, $\times 240$.

further evidence from studies using secondary transfer of granulocyte-macrophage and macrophage colony-forming units (CFU-GM and CFU-M) of AF-1 virus-infected mice (ref. 7 and T.F., unpublished observations). Thus, all of 17 factor-independent CFU-C from bone marrow and all of 6 factor-independent CFU-M from the spleen of DDD mice 5 days after infection with AF-1 virus continued to proliferate after transfer to liquid culture without stimulating factor for 4 months, whereas the proliferation of CFU-M transferred from uninfected controls to cultures without stimulating factors stopped after 4–6 weeks.

Table 1 Effect of AF-1 virus on haematopoietic colony-forming cells in adult DDD mice

			Uninfected	AF-1 virus-infected
Spleen weight (mg)			100	1,787 $\pm$ 180
(mean of three measurements)				
Haematocrit (%)			49	34
CFU-E*	Bone marrow	None	0	0
		+Ep	180 $\pm$ 17	4 $\pm$ 1.0
	Spleen	None	0	0
		+Ep	25 $\pm$ 2.2	15 $\pm$ 1.2
BFU-E	Bone marrow	None	0	0
		+Ep	0	0
		+Wehi	0	0
		+Ep + Wehi	22 $\pm$ 3.1	0
	Spleen	None	0	0
		+Ep	0	0
CFU-C	Bone marrow	+Wehi	0	0
		+Ep + Wehi	3 $\pm$ 0.1	0.3 $\pm$ 0.2
	Spleen	None	0	0
		+Wehi	104 $\pm$ 12	37 $\pm$ 4.2
	Spleen	None	0	20 $\pm$ 1.5
		+Wehi	6 $\pm$ 1.2	22 $\pm$ 0.8

Three-month-old female inbred DDD mice were injected intravenously with 6×10^3 spleen focus-forming units (SFFU) of AF-1 virus and used for the experiment 17 days after infection. AF-1 virus was rescued from 19TC13 non-producer cells with the F-MuLV clone 643/22N. This helper virus alone does not induce spleen focus formation or splenomegaly in adult mice. The technique for the assay of haematopoietic colony-forming cells has been described elsewhere¹⁵. Cells obtained from spleen cell suspensions or flushed out of the femoral cavity were plated at a density of 1×10^5 cells per ml in Iscove's modified Dulbecco's medium (Boehringer, Mannheim) containing 10% fetal calf serum (Flow), 0.8% methylcellulose (Fluka AG, Buchs, Switzerland), 180 μ g iron-saturated human transferrin (Behringwerke, FRG) in 35-mm Petri dishes (Greiner). Where indicated, 0.5 U ml⁻¹ erythropoietin (Ep) partially purified from the urine of patients with Fanconi's anaemia¹⁵, was added. A 10% concentration of Wehi-3 conditioned medium (Wehi) was used as a source of factors (multi-CSF) that stimulate BFU-E and CFU-C formation *in vitro*¹⁶. Erythroid colonies of at least 20 cells were counted as CFU-E 2 days after the establishment of the culture. BFU-E form large erythroid colonies which were counted on day 9 of culture. CFU-C were recorded on the same day as BFU-E and represent the sum of macrophage, granulocytic and macrophage/granulocytic colonies.

* Data are the number of colonies per 10^5 cells plated and were obtained from three independent experiments using three mice.

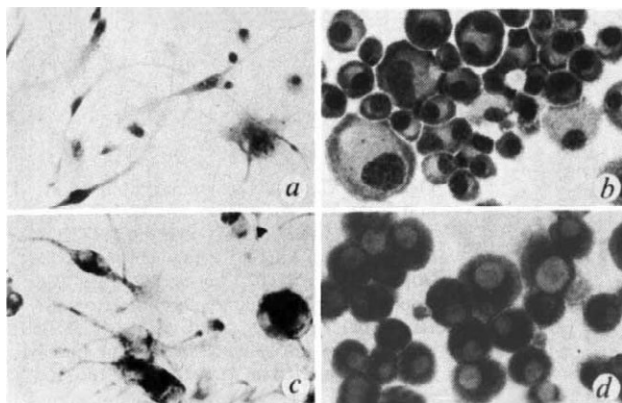


Fig. 3 HA15/A adherent (a, c) and HA15/S non-adherent (b, d) permanent cell lines obtained from spleens of DDD mice after infection with AF-1 virus. a, Cell line HA15/A is characterized by a pleomorphic cell type resembling macrophages. The cells obviously differ in size, cellular configuration and nuclear morphology. The cells are round or bi- and multipolar with round, oval or lobulated nuclei. The pale cytoplasm is heavily vacuolated and possesses long plump or delicately branched processes which occasionally terminate in irregular pseudopods. Multinucleated cells are present (c). b, Cell line HA15/S consists of monoblast-like cells of different size with round, chromatin-dense nuclei, strongly basophilic cytoplasm and prominent paranuclear Golgi bodies. Mono- or multinuclear polyploid cells with atypical nuclear configuration can be recognized. Both cell lines shown strong reactions with acid phosphatase (c, HA15/A cells) and nonspecific esterase (d, HA15/S cells). a, Giemsa, $\times 127$; b, Giemsa, $\times 200$. c, Acid phosphatase using naphthol AS-BI phosphate as a substrate⁶, no nuclear stain $\times 127$. d, Nonspecific esterase using α -naphthyl butyrate as a substrate⁵, no nuclear stain, $\times 200$.

We found that 25% of the factor-independent AF-1 virus-infected CFU-M could be further propagated and maintained as permanent cell lines.

Cell lines were also easily established from spleen cell suspensions of AF-1 virus-infected DDD mice (cell line HA15) as early as 5 days after infection, which suggested early transformation of the target cell by AF-1 virus *in vivo*. The transformed cells proliferated readily in culture requiring neither prior passage of infected tissue *in vivo* nor a long period of tissue culture, unlike, for example, Friend leukaemia cells⁸.

Shortly after establishment, the AF-1 virus-transformed cell lines mainly consisted of adherent slow-growing cells. Forced passage in culture resulted in the isolation of two sublines composed of either mainly adherent cells (HA15/A) or suspended growing cells (HA15/S). Cytologically, HA15/S cells exhibited morphological features of monoblasts and HA15/A cells resembled macrophages (Fig. 3a, b). Both sublines showed activity of nonspecific esterase and acid phosphatase (Fig. 3c, d) but no expression of Mac-1 antigen⁹⁻¹¹. The lack of Mac-1 antigen expression may either be a consequence of the viral transformation or indicate that the AF-1 virus target cell is among the subpopulation of mononuclear phagocytes that does not express this antigen, for example, the histiocytes of the spleen¹². HA15/A and HA15/S cells expressed Fc receptors, whereas only HA15/A cells could phagocytose immunoglobulin-coated polystyrene particles¹³. HA15 cells released virus that transformed fibroblasts and caused spleen focus formation and splenomegaly. The pathogenic property of this virus was not different from the AF-1 virus rescued from the non-producer cells.

Attempts to establish permanent cell lines from spleen cell cultures or CFU-C of uninfected or Ha-MSV-infected mice have failed; the cell lines ceased to proliferate after 2-3 weeks of culture. As both Ha-MSV and the AF-1 virus contain the *ras* oncogene (Fig. 1), it will be of interest to determine whether the immortalization of a subpopulation of mononuclear phagocytes is a property of either the *ras* oncogene or the viral

long terminal repeat of AF-1 virus in combination with the *ras* oncogene. So far, restriction mapping of the genomes of Ha-MSV and AF-1 virus has shown subtle differences between the two in the long terminal repeat region of the genome but not in the *ras* oncogene. Recently, AF-1 virus has been molecularly cloned to pinpoint the genetic elements that make its leukaemogenic properties so unique.

In summary, AF-1 virus is a new murine spleen focus-forming virus carrying a *ras*-related oncogene that possesses the unique property of transforming cells of the mononuclear phagocyte lineage *in vivo*. The AF-1 virus-infected DDD mouse is uniquely appropriate as a model for studying the pathogenesis of human malignant histiocytosis, for which a viral etiology has been proposed¹⁴.

We thank Mrs G. Arman for technical assistance. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (T.F. and W.O.). The Heinrich-Pette-Institut is financially supported by Freie und Hansestadt, Hamburg and by Bundesministerium für Jugend, Familie und Gesundheit, Bonn.

Received 16 July 1984; accepted 6 March 1985.

- Ostertag, W. *et al.* *J. Virol.* **33**, 573-582 (1980).
- Harvey, J. J. *Nature* **204**, 1104-1105 (1964).
- Huhn, D. & Meister, P. *Cancer* **42**, 1341-1349 (1978).
- Rappaport, H. *Tumors of the Hemopoietic System. Atlas of Tumor Pathology* 49-63 (Armed Forces Institute of Pathology, Washington, DC, 1966).
- Vardiman, J. W., Byrne, G. E. & Rappaport, H. *Cancer* **36**, 419-427 (1975).
- Scott, R. B. & Robb-Smith, A. H. T. *Lancet* **ii**, 194-198 (1939).
- Johnson, G. R., Ostertag, W. & Nicola, N. A. in *Modern Trends in Human Leukemia* (ed. Neth, R.) (in the press).
- Ostertag, W. & Pragnell, I. B. *Curr. Topics Microbiol. Immun.* **94/95**, 143-208 (1981).
- Dannenberger, A. M. Jr & Suga, M. in *Methods for Studying Mononuclear Phagocytes* (eds Adams, D. O., Edelson, P. J. & Koren, H. S.) 375-396 (Academic, New York, 1981).
- Leder, L. D. *Der Blutmonozyt* (Springer, Berlin, 1967).
- Springer, T., Galfre, G., Secker, D. S. & Milstein, C. *Eur. J. Immun.* **8**, 539-551 (1978).
- Flotte, T. J., Springer, T. A. & Thorbecke, J. *Am. J. Path.* **111**, 112-124 (1983).
- Taffet, S. M. & Russell, S. W. in *Methods for Studying Mononuclear Phagocytes* (eds Adams, D. O., Edelson, P. J. & Koren, H. S.) 283 (Academic, New York, 1981).
- Lancet* **i**, 455-456 (1983).
- Iscoe, N. N., Guilbert, L. J. & Weyman, C. *Expl. Cell Res.* **126**, 121-126 (1980).
- Iscoe, N. N., Roitsch, C. A., Williams, N. & Guilbert, L. J. *J. cell. Physiol. (Suppl.)* **1**, 65-71 (1982).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Wahl, G. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).
- Ellis, R. W. *et al.* *Nature* **292**, 506-510 (1981).
- Bassin, R. H., Simons, P. J., Chesterman, F. C. & Harvey, J. J. *Int. J. Cancer* **3**, 265-272 (1968).

An HTLV-III peptide produced by recombinant DNA is immunoreactive with sera from patients with AIDS

N. T. Chang, J. Huang, J. Ghayeb, S. McKinney, P. K. Chanda, T. W. Chang, S. Putney*, M. G. Sarngadharan†, F. Wong-Staal† & R. C. Gallo†

Centocor, Inc., 244 Great Valley Parkway, Malvern, Pennsylvania 19355, USA

* Repligen, 101 Binney St, Cambridge, Massachusetts 02139, USA

† Laboratory of Tumor Cell Biology, National Cancer Institute, Bethesda, Maryland 20205, USA

Human T-cell lymphotropic retrovirus type III (HTLV-III), also called lymphadenopathy-associated virus (LAV), has been identified as the aetiological agent of acquired immune deficiency syndrome (AIDS)¹⁻⁹. The sera of most patients with AIDS or AIDS-related complexes, and of asymptomatic individuals infected with HTLV-III, contain antibodies against antigens of HTLV-III^{4,10-12}. The characterization of these antibodies and their corresponding viral antigens is important not only for understanding immunity against HTLV-III and the pathology of AIDS, but also for the development of diagnostic methods and preventive vaccine for AIDS. Following the successful establishment of a long-term T-cell line permissive for HTLV-III replication¹, large quantities of virus have been produced, facilitating the purification of viral proteins and the development of mouse monoclonal antibodies

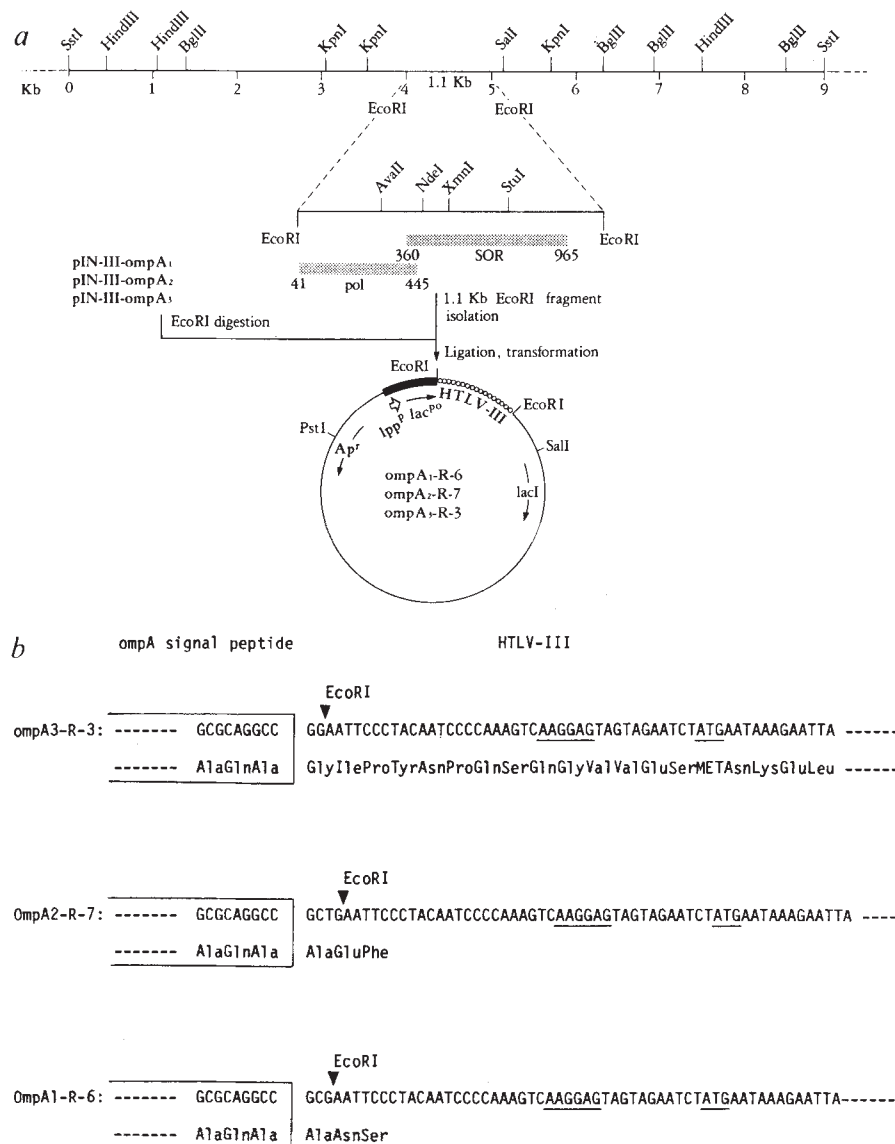


Fig. 1 a, Construction of recombinant plasmids ompA1-R-6 (01R6), ompA2-R-7 (02R7) and ompA3-R-3 (03R3) carrying HTLV-III inserts. The 1.1-kb *EcoRI* fragment was excised from clone λ BH10, purified by gel electrophoresis and ligated to plasmid vectors pIN-III-ompA1, -ompA2 and -ompA3 which had been cut previously with *EcoRI*. The ligation mixture was used to transform *E. coli* K-12 strain JA221 (ref. 17) and the resultant ampicillin-resistant colonies were isolated for further characterization. The 15,000- M_r peptide is presumed to start at nucleotide 41 and end at position 445 and the sequence of the short open reading frame extends from nucleotides 360 to 965 (see text). The top line shows the restriction map of HTLV-III insert of clone λ BH10 (refs 13, 16). The dotted line indicates the left-hand and right-hand arms of λ vector λ gtWES λ B. The solid line indicates the HTLV-III DNA. The two shaded regions indicate the two open reading frames included in the 1.1 kb *EcoRI* fragment, the 3' end of the *pol* gene and the SOR. **b**, The nucleotide sequence of the ompA signal peptide and the pertinent region of recombinant plasmids 01R6, 02R7 and 03R3. The sequence of the ompA signal peptide is boxed. The postulated Shine-Dalgarno sequence AAGGAG and the initiation codon ATG are underlined. Abbreviations: *lpp*^P, *lpp* promoter, *lac*^{PO}, *lac* promoter-operator; AP^r, ampicillin-resistance gene.

against several viral antigens^{10,24}. More recently, the structure of HTLV-III proviral DNA has been elucidated¹³⁻¹⁶. We now report the production, by genetic engineering methods, of a peptide encoded by a gene segment of HTLV-III. A 1.1-kilobase (kb) *EcoRI* DNA segment from an isolate of HTLV-III was inserted into a *lpp* and *lac* promoter-coupled expression vector, pIN-III-ompA. *Escherichia coli* transformants of this plasmid produced a peptide of relative molecular mass (M_r) 15,000 (15K) which was strongly immunoreactive with anti-HTLV-III antibodies present in sera from AIDS patients. Lysates of the clones expressing this 15K peptide inhibited the reactivity of the p31 virion protein with AIDS sera, suggesting that it is a fragment of the viral p31 protein. The peptide reacted with sera from all 20 AIDS patients but none of the 8 normal controls tested. These results suggest that the peptide may be useful for detecting anti-HTLV-III antibodies in blood samples.

The expression vector used, pIN-III-ompA (ompA), contains the lipoprotein (the most abundant protein in *E. coli*) gene promoter (*lpp*) and the *lacUV5* promoter-operator (Fig. 1a)¹⁷. The ompA vectors also contain the DNA segment encoding the lac repressor, which allows the expression of the inserted DNA to be regulated by *lac* operon inducers such as isopropyl- β -D-thiogalactoside (IPTG). The ompA cloning vehicles contain three unique restriction enzyme sites, *EcoRI*, *HindIII* and *BamHI*, in all three reading frames, so that DNA can be inserted into any of these restriction sites.

Various restriction fragments were excised from the recombinant clone λ BH10, which contains a 9-kb HTLV-III DNA

insert in the *SstI* site of the vector λ gtWES λ B (ref. 13). These restriction fragments were then inserted into the ompA vectors in all three reading frames and used to transform *E. coli* JA221 cells¹⁷. Transformants were first screened for HTLV-III DNA by *in situ* colony hybridization using HTLV-III DNA probes¹⁸, and the positive clones then screened for expression of HTLV-III antigenic peptides using HTLV-III-specific antibodies.

This approach identified several gene segments that encode peptides showing immunoreactivity with anti-HTLV-III antibodies, among these a 1.1-kb *EcoRI* restriction fragment. When this fragment was inserted into the ompA vectors in the three reading frames (Fig. 1a), IPTG induction of transformants of all three plasmid constructs, designated ompA1-R-6 (01R6), ompA2-R-7 (02R7) and ompA3-R-3 (03R3), produced a 15K peptide that was strongly reactive with anti-HTLV-III antibodies in sera from AIDS patients (Fig. 2). This reactivity was not detected with sera from normal individuals (data not shown).

DNA sequence analysis of the three recombinant constructs confirmed that, as planned, each recombinant had a different reading frame of the HTLV-III (+)-strand coupled to the coding sequence of each vector. Only in 03R3 was the reading frame of the inserted DNA in phase with that set by the signal peptide in the ompA vector; in 01R6 and 02R7 the *pol* gene segment DNA was out of phase (Fig. 1b). Examining the DNA sequence closely, we found that there was a 6-base pair (bp) ribosome binding site, AAGGAG (Shine-Dalgarno sequence), located at nucleotide position 24-29 and an initiation codon, ATG, located 11 bp downstream (position 41-43). We believe that the 15K

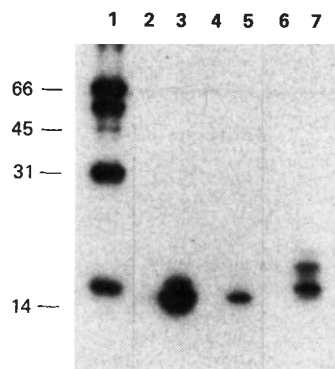


Fig. 2 Immunoreactivity with AIDS sera of HTLV-III antigens produced in bacterial cells containing plasmids 01R6, 02R7 or 03R3. Lane 1, purified HTLV-III virions; lanes 2, 3, 01R6 uninduced and induced with IPTG; lanes 4, 5, 02R7 uninduced and induced; lanes 6, 7, 03R3 uninduced and induced. Numbers on the right represent M_r standards ($\times 10^{-3}$).

Methods. Cells were grown at 37°C in L-broth containing 100 $\mu\text{g ml}^{-1}$ ampicillin to an A_{600} of 0.2. The cell cultures were then divided into two aliquots; IPTG was added to one aliquot to a final concentration of 2 mM (induced) and the other aliquot received no IPTG (uninduced). After 2 h incubation at 37°C, the cells were collected and lysed in gel loading buffer²². Total cell lysates from the equivalent of 0.2 ml cell culture or purified NP40-disrupted HTLV-III virions were electrophoresed on 12.5% SDS-polyacrylamide gels by the method of Laemmli²². The proteins on the gels were then electrophoretically blotted onto a 0.45- μm nitrocellulose filter²³. The blots were first incubated with a blocking buffer (5% non-fat dry milk, 0.1% Antifoam A, 0.1% Na-azide in phosphate-buffered saline) for 2 h at 37°C and then with fresh blocking buffer containing 5% goat serum for 2 h at room temperature. AIDS patients' or normal donor sera were then added to a final dilution of 1:100 and the blots were incubated at 4°C overnight. After incubation the filters were washed three times with washing buffer (0.5% deoxycholic acid, 0.1 M NaCl, 0.5% Triton X-100, 10 mM phosphate pH 7.5, 0.1 mM phenylmethylsulphonyl fluoride) for 30 min each and then incubated with fresh blocking buffer containing 5% normal goat serum for 1 h before adding ¹²⁵I-labelled goat anti-human IgG ($10 \mu\text{Ci } \mu\text{g}^{-1}$) to a final dilution of 1×10^6 c.p.m. ml^{-1} . After 30 min, the iodinated probe was removed and the nitrocellulose blots were washed with washing buffer three times for 20 min each, dried and exposed to Kodak XAR-5 X-ray film.

peptide synthesized by all three recombinants is translated from the transcripts using this internal initiation codon. If this is true, the peptide would start from the ATG located at position 41–43 and end at the stop codon at position 446–448, producing a peptide of 135 amino-acid residues encoded by the 3'-end segment of the *pol* gene of HTLV-III. As further evidence that the 15K peptide is indeed derived from the *pol* gene, clones containing only the 5'-end *EcoRI/NdeI* fragment (Fig. 1a) still produced the same 15K peptide (data not shown).

In addition to the 15K peptide, the 03R3 construct, in which the reading frame of the HTLV-III DNA *pol* gene is in phase with that set by the vector, produced another two peptides ~19K and 16.5K in size (Fig. 2). It is possible that the 19K peptide contains an additional 35 amino-acid residues, 21 of which are from the signal peptide encoded by the *ompA3* vector, with 14 encoded by the inserted HTLV-III DNA itself. The 16.5K peptide may be the processed 19K peptide in which the signal peptide is cleaved.

The 01R6 and 02R7 constructs also produced another peptide of ~17.5 K (Fig. 2) which was weakly reactive with sera from AIDS patients, but whose origin is unclear. Note that the 1.1-kb *EcoRI* fragment contains a second potential coding region, designated as the short open reading frame (SOR), extending from nucleotides 360 to 965 (Fig. 1a)¹⁶. Four of the five AUG methionine codons in this region are near the 5' end of this open reading frame. However, there is no clearly recognizable ribosome binding site at the 5' end of this open reading frame.

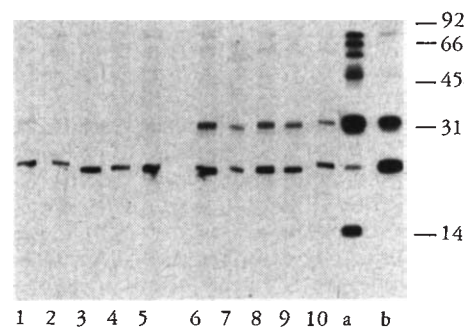


Fig. 3 Inhibition of reaction between HTLV-III antigens and an AIDS serum by 01R6 lysate. Lanes 1–5, 01R6 cell lysate of 50, 75, 100, 200 and 400 μl , respectively; lanes 6–10, control cell lysate of 50, 75, 100, 200 and 400 μl , respectively; lane a, no cell lysate (with serum from AIDS patient A60); lane b, no cell lysate (with serum from AIDS patient E.Y.).

Methods. Clone 01R6 or control cells in 10-ml cultures were induced with IPTG at an A_{600} of 0.2 and the cells were collected after 2 h, resuspended in 1 ml 10 mM Tris-HCl pH 8.0, and sonicated. NP40-treated HTLV-III was fractionated on a 12.5% SDS-polyacrylamide gel and electro-transblotted onto a nitrocellulose filter paper. The paper was cut into ~30 strips which were then incubated with a well-characterized AIDS serum (from patient E.Y., which had elevated anti-p31 antibodies) in the presence of increasing amounts of lysate of clone 01R6 (lanes 1–5) or a negative control (lanes 6–10). The lysates were then used in the competition inhibition immunoassay. The further immunoblotting procedure was performed as described for Fig. 2.

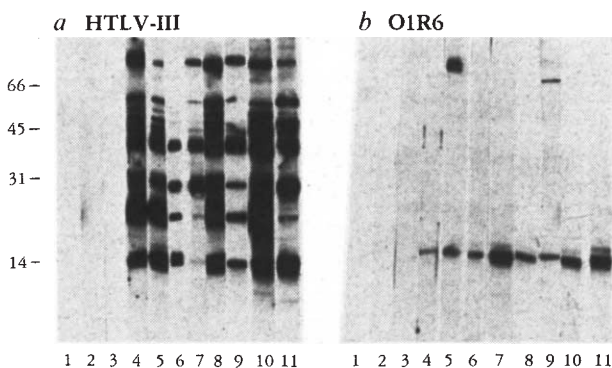


Fig. 4 The presence of antibodies against the recombinant HTLV-III-encoded 15K peptide in sera from AIDS patients and healthy individuals. Purified HTLV-III virus (a) or total cell lysate of bacterial clone 01R6 (b) were electrophoresed and electroblotted onto nitrocellulose filters, and protein strips were reacted with serum samples as described for Fig. 3. Lanes 1–3, healthy donors; lanes 4–11, AIDS patients.

Furthermore, several recombinant clones containing DNA fragments of the SOR, when inserted into the open reading frame vector pMR100, produced λ CI-HTLV-III- β -galactosidase tripartite fusion proteins which showed little immunoreactivity with anti-HTLV-III antibodies in sera from AIDS patients¹⁹.

Because the RNA-dependent DNA polymerase of retroviruses is a minor component of the virions, we were surprised to detect immunoreactivity against the 15K peptide in sera from AIDS patients. To identify this immunoreactive peptide with respect to the banding pattern of HTLV-III virion antigen in SDS-polyacrylamide gel electrophoresis, we used a competition inhibition immunoassay. Identical filter strips containing electrophoresed disrupted HTLV-III virions were incubated with a well-characterized serum from an AIDS patient in the presence or absence of lysates of 01R6, or control bacterial clones. Lysates of 01R6 blocked the immunoreactivity of the viral p31 protein with the AIDS serum, but lysates of control cells did not (Fig. 3). This result suggests that the recombinant 15K peptide encoded by the 3' end of the viral *pol* gene is also a part of another virion protein, p31, in contrast to a previous proposal

that p31 is a cellular protein which co-purifies with HTLV-III virions¹⁰.

We next evaluated the prevalence in AIDS patient sera of antibodies against the 15K peptide. A panel of sera from AIDS patients and normal healthy individuals was tested in Western blot analyses using 01R6 lysate as the source of antigen. All of 20 AIDS sera and none of the 8 normal controls reacted with the 15K peptide (Fig. 4), indicating that most, if not all, AIDS patients produce antibodies against the viral p31 protein.

Thus, the 15K peptide described here seems to be encoded by the 3' end of the viral *pol* gene located at the 5' side of the 1.1-kb *EcoRI* fragment and to be related to the HTLV-III virion protein p31. The precise nature of the viral p31 protein remains unclear. However, in other retroviruses, such as avian myoblastosis virus and murine leukaemia virus, the 3' end of the *pol* gene has been shown to encode, in addition to the reverse transcriptase, an endonuclease, RNase H, which seems to be essential for virus replication^{20,21}. The p31 peptide encoded by the 3' end of the HTLV-III *pol* gene may represent the corresponding RNase H of this virus. Work is in progress to purify and sequence the 15K peptide from bacterial clones expressing it. The peptide is also being evaluated for use as the solid-phase immunoadsorbent in immunoassays screening for anti-HTLV-III antibodies in serum samples.

We thank Dr M. Inouye for providing the ompA vectors, Dr R. Ting (Biotech) for the purified HTLV-III, Dr V. R. Zurawski Jr for helpful discussions, Dr H. Schoemaker and Mr M. Wall for supporting the work, and S. O'Haire and D. Scollon for typing the manuscript.

Received 20 December 1984; accepted 19 February 1985.

1. Popovic, M., Sarngadharan, M. G., Read, E. & Gallo, R. C. *Science* **224**, 497-500 (1984).
2. Gallo, R. C. *et al. Science* **224**, 500-503 (1984).
3. Schüpbach, J. *et al. Science* **224**, 503-505 (1984).
4. Sarngadharan, M. G., Popovic, M., Bruch, L., Schüpbach, J. & Gallo, R. C. *Science* **224**, 506-508 (1984).
5. Vilmer, E. *et al. Lancet* **i**, 753-757 (1984).
6. Groopman, J. E. *et al. New Engl. J. Med.* **311**, 1419-1422 (1984).
7. Barre-Sinoussi, F. *et al. Science* **220**, 868-870 (1983).
8. Broder, S. & Gallo, R. C. *New Engl. J. Med.* **311**, 1292-1297 (1984).
9. Markham, P. D., Shaw, G. M. & Gallo, R. C. in *AIDS* (eds Devita, V. T., Helman, S. & Rosenberg, S. A.) (Lippincott, Philadelphia, in the press).
10. Safai, B. *et al. Lancet* **i**, 1438-1440 (1984).
11. Kitchen, L. W. *et al. Nature* **312**, 367-369 (1984).
12. Cheingsong-Popov, R. *et al. Lancet* **ii**, 477-480 (1984).
13. Hahn, B. H. *et al. Nature* **312**, 166-169 (1984).
14. Ayra, S. K. *et al. Science* **225**, 927-929 (1984).
15. Shaw, G. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 4544-4548 (1984).
16. Ratner, L. *et al. Nature* **313**, 277-284 (1985).
17. Ghayeb, J. *et al. EMBO J.* **3**, 2437-2442 (1984).
18. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237-251 (1977).
19. Chang, N. T. *et al. Science* **228**, 93-96 (1985).
20. Mölling, K. *et al. Nature new Biol.* **234**, 240-243 (1971).
21. Levin, J. G., Hu, S. T., Rein, A., Messer, L. I. & Gerwin, B. I. *J. Virol.* **51**, 470-478 (1984).
22. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
23. Towbin, H., Staehlin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350-4354 (1979).
24. Veronese, F. *et al. Proc. Natn. Acad. Sci. U.S.A.* (in the press).

Table 1 Interaction of HBV (HBsAg) with human hepatoma cells

Cells used	Antibodies added	Percentage of cells attached to HBsAg-cellulose
HepG2 (human hepatoma)	—	80-95 (14 experiments)
HeLa	—	10-20
Rat hepatocytes	—	10-20
HepG2	Anti-pre-S(120-145) (15 mg IgG ml ⁻¹)	60
HepG2	Anti-pre-S(12-32) (15 mg IgG ml ⁻¹)	30
HepG2	Anti-pre-S(120-145) + anti-pre-S(12-32) (7.5 mg IgG ml ⁻¹ each)	20
HepG2	Anti-HBV serum devoid of anti-S-protein ⁵	8
HepG2	—	0 (BSA-cellulose)

HBV (HBsAg) was attached to *N-N'*-*p*-phenylenedimaleimide-derivatized sulphhydryl cellulose under conditions described for linking of pre-S(120-145)². About 4 mg of HBV (HBsAg) was linked to 1 g of the cellulose derivative. A control cellulose derivative was prepared by linking bovine serum albumin (BSA) to the activated matrix. The cellulose derivative (40 mg) suspended in TS (0.14 M NaCl, 0.01 Tris-HCl, 0.02% NaN₃, pH 7.2) containing 10 mg ml⁻¹ BSA (TS-BSA) was mixed with ~2 × 10⁶ washed HepG2 human hepatoma cells¹⁵, suspended in TS-BSA and incubated for 30 min at 37 °C, followed by 1 h at 4 °C. HeLa cells and clone 9 normal rat liver cells (American Type Culture Collection) were used as controls. The cell/cellulose mixtures were layered on top of 1 ml of 33% (w/w) Hypaque and centrifuged at 3,000 r.p.m. for 2 min. The cellulose derivative with attached cells pelleted in these conditions; unattached cells recovered from the Hypaque/TS-BSA interphase were diluted five-fold in TS-BSA and pelleted by centrifugation. The relative proportion of adsorbed and unadsorbed cells was determined by measurement of lactate dehydrogenase (LDH) activity in appropriate aliquots of cell lysates obtained after exposure to the detergent Triton X-100 (5 mg ml⁻¹ in H₂O). LDH activity was determined using Sigma diagnostic kit no. 500. In initial experiments, similar results were obtained with HBV-cellulose and HBsAg-cellulose. Because of the limited quantities of HBV available, HBsAg-cellulose was used in subsequent experiments.

protein corresponding to the entire *env* gene (pre-S + S). Synthetic peptides from the N-terminals of proteins (2) and (3), and antisera to them have been used to study the occurrence and properties of pre-S sequences. The results presented here provide unambiguous evidence that all three *env* encoded proteins are present in HBV particles; synthetic peptides corresponding to the gene encoding pre-S are highly immunogenic and can be used in diagnostic tests for detection in human sera of antibodies preferentially recognizing HBV; such antibodies, specific for pre-S determinants, are elicited during hepatitis B infection and by immunization with HBV proteins (2) and (3); the hepatitis B vaccine licensed in the United States does not contain pre-S proteins; and the pre-S proteins of the HBV envelope contain domains specifically recognized by liver cells. These findings suggest that pre-S determinants are important in virus-neutralizing responses and should be present in HBV vaccines.

Peptides corresponding to the N-terminals of proteins (2) and (3) were synthesized and used to raise antisera in rabbits. High levels of anti-peptide antibodies were elicited (Fig. 1A, B).

The use of anti-pre-S (120-145) in Western blots of HBV particles identified components of relative molecular mass (*M_r*) ~33,000 (~33K), ~36K, ~45K and ~67K (Fig. 1C, lane d). A similar experiment with anti-pre-S (12-32) identified the ~45K protein alone (data not shown), indicating that this component corresponds to protein (3), the full-length translation product of the *env* gene, while the ~33K and ~36K components correspond to protein (2)²⁻⁴. The ~45K component was not detectable in Western blots of 22-nm sub-viral particles lacking the nucleocapsid (hepatitis B surface antigen, HBsAg)². This is consistent with previous evidence for antigenic determinants on the HBV envelope which are underrepresented or absent from

Hepatitis B virus contains pre-S gene-encoded domains

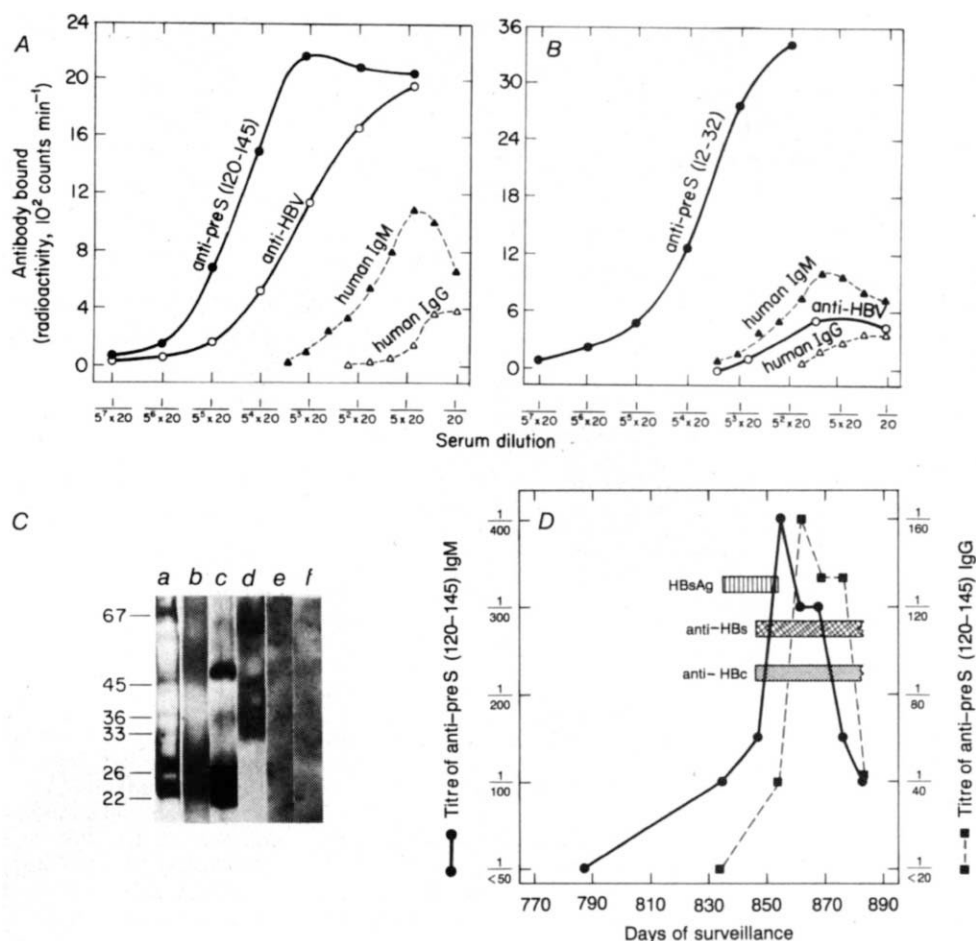
A. Robert Neurath*, Stephen B. H. Kent†, Nathan Strick*, Patricia Taylor* & Cladd E. Stevens*

* The Lindsley F. Kimball Research Institute of The New York Blood Center, New York, New York 10021, USA

† Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

One of the open reading frames on hepatitis B virus (HBV) DNA comprises the coding region (designated the *env* gene) for the virus envelope proteins. Studies on messenger RNA transcription suggest that this gene has the potential to code for three related proteins¹: (1) a protein of 226 amino acids identified as a major protein constituent of the HBV envelope, termed S-protein; (2) a protein with 55 additional amino acids at the N-terminal coded for by a portion of the *env* gene upstream of the S-gene (pre-S); (3) a

Fig. 1 A, B. Radioimmunoassays of IgG antibodies in serial dilutions of: rabbit antisera to pre-S (120-145) (A) or pre-S (12-32) (B) (●), to HBV (○) (devoid of antibodies to S-protein detectable by RIA⁵); IgG (Δ) and IgM (▲) in serum of a patient who had recovered from hepatitis B. The latter serum was obtained before antibodies to the S-protein were detectable. Immulon 2 Removawell strips (Dynatech Laboratories) were coated with 20 μg ml⁻¹ of either free pre-S (120-145) or pre-S (12-32), then coated with gelatine (2.5 mg ml⁻¹ in 0.1 M Tris, pH 8.8). The conditions for coating and the double-antibody RIA were described elsewhere^{2,13}. The peptides corresponding to subtype adw2 (ref. 16) were synthesized as described elsewhere²; their sequences are MQWNSTAFHQTLQDPRYRGLYLPAGG [pre-S (120-145)] and MGTNLSVPNPLGFPPDHQLDP [pre-S (12-32)], according to the one-letter amino acid code. **C.** Polyacrylamide gel electrophoresis of HBV (a, d), serum-derived HBsAg after pepsin treatment (components of the Merck, Sharp & Dohme hepatitis B vaccine)¹⁰ (b, e), and HBsAg produced in yeast¹¹ and lacking sequences corresponding to pre-S protein (c, f). Separated polypeptides were either stained with silver (a-c) or transferred to nitrocellulose and reacted with anti-pre-S (120-145) followed by ¹²⁵I-labelled protein A (d-f)². **D.** Development of antibodies to the pre-S gene-encoded protein of HBV during acute hepatitis B infection. Antibodies to pre-S (120-145) were quantified; HBsAg, anti-HBs and anti-HBc were assayed using commercial test kits (Abbott Laboratories). The broken line at the end of the bars corresponding to the different markers of HBV infection indicates their presence at the end of surveillance. Antibody titres represent the highest dilution of serum at which radioactivity counts corresponding to the specimens, divided by counts corresponding to equally diluted control serum, were ≥2.1.



HBsAg⁵⁻⁹. To confirm this, we studied the reaction of an anti-HBV serum with two synthetic peptide analogues of the pre-S protein; the antiserum reacted at high dilutions with both peptides (Fig. 1A, B).

The synthetic peptides were also recognized by antibodies in sera of individuals who had just recovered from acute hepatitis B (Fig. 1A, B). Figure 1D shows the time course of development of antibodies recognizing pre-S (120-145) in a selected patient; similar results were obtained when antibodies recognizing pre-S (12-32) were assayed (data not shown). IgM antibodies recognizing the peptides were detected during HBsAg antigenaemia before antibodies to the S-protein (anti-HBs) or to hepatitis B core antigen (anti-HBc) were detectable. After production of the latter two antibodies, the titre of antibodies with anti-pre-S specificity declined. In some cases, antibodies recognizing the synthetic peptides were present even before HBsAg was detected in plasma, or when HBsAg never appeared in blood and the only marker for hepatitis B was anti-HBc and later anti-HBs.

HBV particles were recognized much more efficiently than HBsAg in radioimmunoassay (RIA) tests with anti-pre-S (120-145) (Fig. 2) or with antibodies to HBV particles (data not shown); this result implies that determinants specific for the preS region of protein (2) are more abundant in HBV than in HBsAg. Treatment of HBsAg with pepsin, a step used in the preparation of hepatitis B vaccine¹⁰, resulted in a ~10³-fold decrease in reactivity to anti-pre-S (120-145) (Fig. 2). HBsAg from vaccines derived either from infected plasma or produced in yeast¹¹ had ≤1/5,000 of the reactivity of intact HBsAg in these tests (data not shown; see also Fig. 1C). Human subjects vaccinated with either of these vaccines did not develop detect-

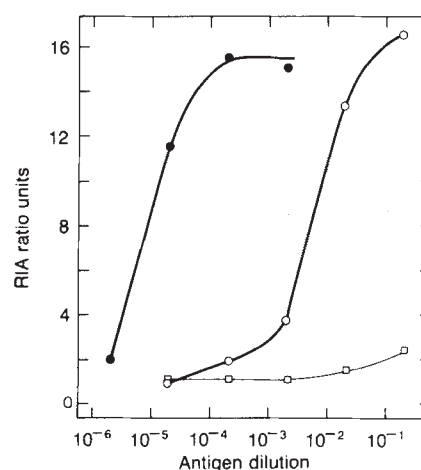


Fig. 2 Radioimmunoassays of various preparations containing HBV-specific proteins on polystyrene beads coated with anti-pre-S (120-145) IgG. The antigens tested were: HBV (●), HBsAg (○) and HBsAg treated with pepsin (1 mg ml⁻¹ HBsAg, 50 μg ml⁻¹ pepsin in 0.1 M glycine-HCl pH 2.2, 2 h at 37 °C) (□). RIA performed as described elsewhere². The concentration of HBsAg S-protein was adjusted to the same level in all preparations based on RIA tests (AUSRIA, Abbott Laboratories).

able antibodies recognizing either of the two synthetic peptides, whereas 7 out of 12 individuals who received a vaccine consisting of intact HBsAg¹² developed these antibodies.

Quantitative aspects of the immunological cross-reactivity between pre-S gene-encoded sequences exposed on HBV

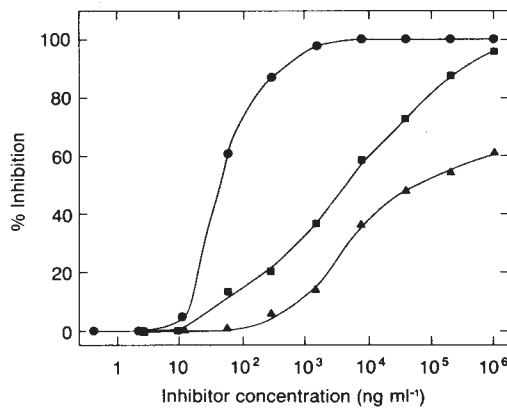


Fig. 3 Inhibition of the reaction of anti-pre-S (120-145) IgG with a pre-S(120-145)- β -galactosidase conjugate. $\bullet$, pre-S(120-145). The HBV ($\blacksquare$) and HBsAg ($\blacktriangle$) preparations contained the same concentration of HBsAg S-protein as determined by RIA (AUSRIA, Abbott Laboratories). The inhibition tests and the preparation of peptide- β -galactosidase conjugates were as described elsewhere^{13,17}. Anti-pre-S(120-145) did not react with unconjugated β -galactosidase. The concentration of free peptide sufficient for ~50% inhibition of the reaction of pre-S(120-145)- β -galactosidase with anti-pre-S(120-145) is ~1/100 (w/w) of that for HBV. However, as the M_r of pre-S(120-145) is ~33K and that of HBV protein components reacting with anti-pre-S(120-145) (representing a minor (<20%) portion of the total HBV mass) is between ~33K and ~67K, the calculated molar concentration of the HBV pre-S protein required for 50% inhibition is approximately the same as that of the peptide analogue.

particles (or on HBsAg) and the synthetic peptide analogues were studied in more detail. The results (Fig. 3) indicate that HBV inhibited completely the reaction between anti-pre-S (120-145) and pre-S (120-145)- β -galactosidase. There is no significant population of anti-peptide antibodies which recognizes the synthetic peptide but not the native protein. Such antibody subpopulations are frequently observed in other antisera raised against synthetic peptide analogues of viral proteins^{13,14}. The molar concentrations of HBV and pre-S (120-145) required for 50% inhibition were approximately the same, indicating that the antigenic determinants on the peptide analogue and on the corresponding segment of the HBV envelope protein(s) are identical or closely related. HBsAg has about one-fifth of the inhibitory activity of HBV. The inhibitory activity of pepsin-treated HBsAg was $\leq 1/1,000$ of the activity corresponding to intact HBsAg.

To explore the involvement of pre-S gene-encoded HBV domains in attachment to cell receptors, we studied the reaction of HBsAg and HBV with liver cells. The results (summarized in Table 1) show that HBsAg (HBV) reacts with human hepatoma cells and that this reaction is inhibited by antibodies to synthetic peptides corresponding to the pre-S gene. Normal rabbit IgG, as well as antibodies to the S-protein (elicited by immunization with pepsin-treated HBsAg), had no inhibitory effects.

In conclusion, pre-S determinants clearly play an important role in hepatitis B virus infection and in the human immune response to HBV, and therefore should be included in hepatitis B vaccines.

We thank T. Huima, P. Sproul and E. Seda for technical assistance; Dr D. G. George for computer searches; Drs E. Emini, W. J. McAleer, F. B. Hollinger and J. L. Dienstag for hepatitis B vaccines and human sera; and Dr W. J. Rutter for providing us with unpublished data.

Received 10 October 1984; accepted 7 March 1985.

1. Rutter, W. J. *et al.* in *Viral Hepatitis and Liver Disease* (eds Vyas, G. N., Dienstag, J. L. & Hoofnagle, J.) 67-86 (Grune & Stratton, Orlando, Florida, 1984).
2. Neurath, A. R., Kent, S. B. H. & Strick, N. *Science* **224**, 392-395 (1984).
3. Stibbe, W. & Gerlich, W. H. *J. Virol.* **46**, 626-628 (1983).
4. Machida, A. *et al.* *Gastroenterology* **86**, 910-918 (1984).
5. Neurath, A. R., Trepo, C., Chen, M. & Prince, A. M. *J. gen. Virol.* **30**, 277-285 (1976).
6. Moodie, J., Stannard, L. M. & Kipps, A. *J. gen. Virol.* **24**, 375-379 (1974).

7. Neurath, A. R., Strick, N. & Huang, C. Y. *J. gen. Virol.* **48**, 53-61 (1980).
8. Alberti, A., Diana, S., Sculard, G. H., Eddleston, A. L. W. F. & Williams, R. *Br. med. J.*, 1056-1058 (1978).
9. Heermann, K. H. *et al. J. Virol.* **52**, 396-402 (1984).
10. Hilleman, M. R. *et al.* in *Viral Hepatitis, 1981 int. Symp.* (eds Szmuness, W., Alter, H. J. & Maynard, J. E.) 385-397 (Franklin Institute Press, Philadelphia, 1982).
11. McAleer, W. J. *et al. Nature* **307**, 178-180 (1984).
12. McAuliffe, V. J., Purcell, R. H., Gerin, J. L. & Tyeryar, F. J. in *Viral Hepatitis* (eds Szmuness, W., Alter, H. J. & Maynard, J. E.) 425-435 (Franklin Institute Press, Philadelphia, 1982).
13. Neurath, A. R., Kent, S. B. H. & Strick, N. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7871-7875 (1982).
14. Komoriya, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 1351-1355 (1984).
15. Aden, D. P., Fogel, A., Plotkin, S., Damjanov, I. & Knowles, B. B. *Nature* **282**, 615-617 (1979).
16. Valenzuela, P. *et al. Nature* **280**, 815-819 (1979).
17. Neurath, A. R., Kent, S. B. H. & Strick, N. *Virus Res.* **1**, 321-331 (1984).

Intragenic recombination leads to pilus antigenic variation in *Neisseria gonorrhoeae*

Per Hagblom*†, Ellyn Segal†, Elizabeth Billyard† & Magdalene So†

* Department of Microbiology, University of Umeå, S-901 87 Umeå, Sweden

† Department of Molecular Biology, Scripps Clinic and Research Foundation, 10666 N. Torrey Pines Rd, La Jolla, California 92037, USA

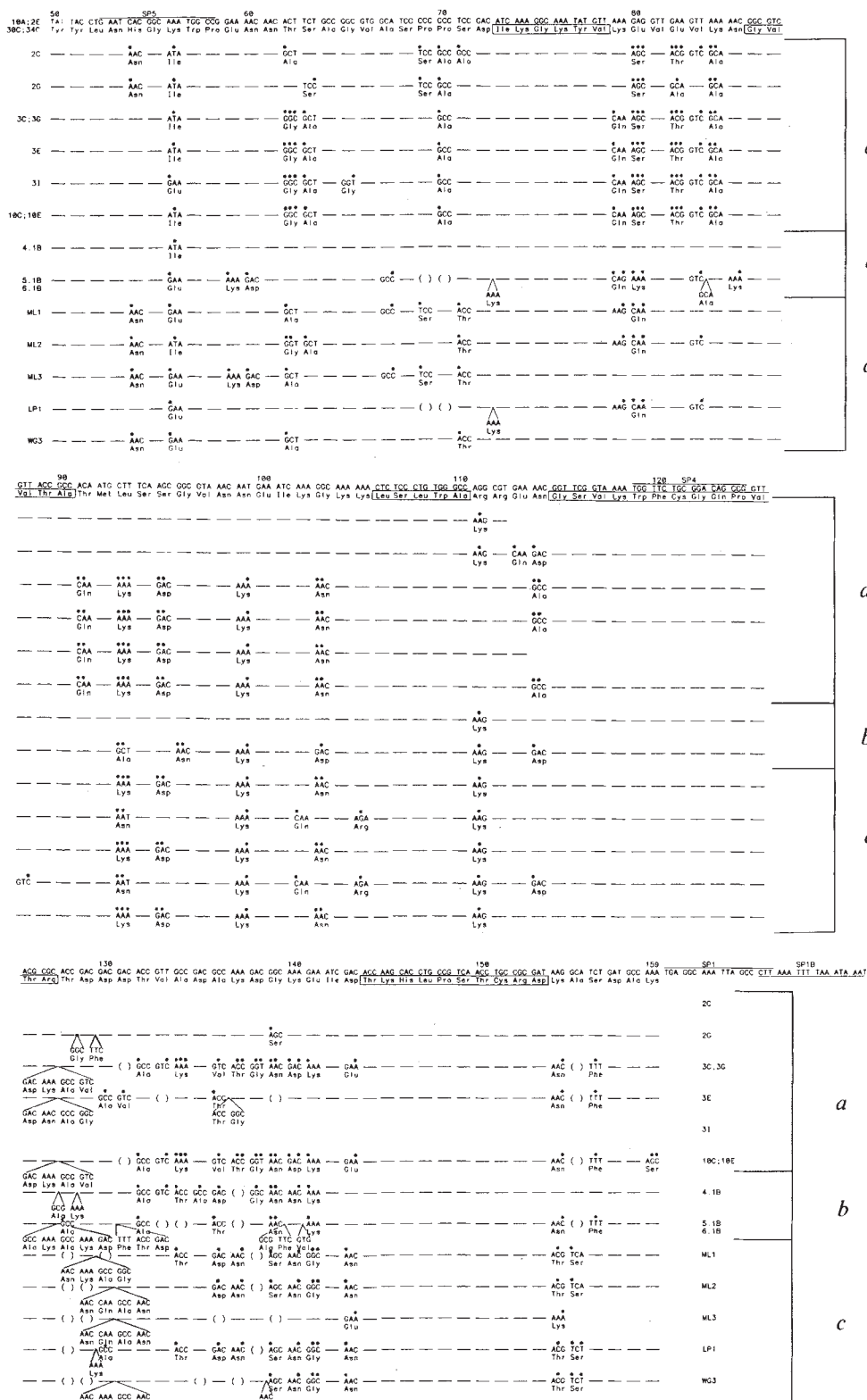
The pilus of the bacterium *Neisseria gonorrhoeae* is a fimbriate surface structure which promotes attachment of the bacterium to host epithelial cells¹⁻⁴. Gonococcal pilus phase variation is characterized by a rapid on/off switch in which pilliated (P^+) cells throw off non-pilliated (P^-) variants and vice versa. Two regions of the gonococcal chromosome (*pilE1* and *pilE2*) act as pilin expression loci⁵, reminiscent of the *MAT* locus in the yeast *Saccharomyces cerevisiae*⁶, while several other chromosomal regions contain silent (non-expressing) pilin sequences⁵. Biochemical and antigenic diversity is seen in pili from a wide variety of clinical isolates⁷⁻¹¹. Pilins (pilin subunits) are composed of conserved N-terminal and variable C-terminal regions¹²; the conserved region of gonococcal pilin is also found in pilins produced by widely disparate bacteria¹³⁻¹⁵. We show here that the gonococcal pilin undergoes antigenic variation *in vitro* and *in vivo*. The protein consists of constant, semi-variable and hypervariable regions. This antigenic variation probably involves gene conversion of mini-cassettes of pilin information.

To study pilus antigenic variation in *N. gonorrhoeae*, lines of phase variants of strain MS11 were established whose members differ from each other in their piliation state (Fig. 2). Pilin transcripts from these derivatives were sequenced by the primer extension method¹⁶. MS11 can express at least seven different pilin genes (Figs 1a, 2). Within one line of derivatives, the P^- to P^+ switch resulted in the expression of a new variant (for example, 2C, E, G) or the re-expression of a former variant (for example, 3G, 10E). Thus, pilin information is not lost during phase or antigenic switching. Those derivatives which express variant genes produce pili that are antigenically different from those of the MS11 progenitor (Fig. 2).

Attempts to obtain sequences of the 3'-untranslated region of the pilin messenger RNA of variants 2C and 3I were unsuccessful (Fig. 1). These alternative sequences are probably the result of aberrant recombination events occurring between a silent locus and the expression site. Our sequence data on several expression loci in the P^- state indicate that this region is often affected by the events regulating the P^+ to P^- switch¹⁷.

The pilin protein can be divided into three regions (Fig. 1). The first 53 amino acids are always conserved. A silent change has occurred at position 47 in variants 2C, ML3 and WG3 (data not shown), suggesting that the constant (C) region may be shorter than 53 residues. Many amino-acid substitutions are found within the semi-variable (SV) region (residues 54-114). Most changes here involve charged amino acids, but the introduction of a positive charge is often balanced by a second

Fig. 1 Pilin gene sequences including deduced amino-acid sequences from residue 50 to the C-terminus. The MS11 pilin DNA and amino-acid sequences⁵ are shown at the top. This sequence is also found in variants 2E, 10A, 30C and 34C. *a*, Sequences from the MS11-derived P⁺ strains; *b*, pilin gene sequences in the expression loci of the MS11-derived P⁻ variants; *c*, sequences of the pilins from the clinical isolates. Strains ML1, 2 and 3 were obtained from the same individual at different times; LP1 and WG3 were isolated from two other individuals (P. F. Sparling, manuscript in preparation). Dashes indicate no sequence differences observed in these positions. Parentheses indicate codon deletions. Codon and amino-acid changes are indicated where found. If the sequence change does not alter the amino acid at that position, only the codon change is shown. Dots above the bases indicate those positions within the codon which have changed. Insertions are presented below the variant sequences. Boxed amino acids in the top line represent the regions found to be conserved by comparison of all variant sequences. Primers were made by an Applied Biosystems 380A oligonucleotide synthesizer to regions which, by comparison of the different sequences, were observed to be invariant. Primer extension sequencing was done essentially as described previously¹⁶, except that kinase-treated primers were used. The reverse complement of the sequence of each primer used is indicated by a line above or below the topmost DNA sequence. As a control for the accuracy of the technique, we have sequenced the 10A derivative and found it to be identical to that of the pilin gene from the progenitor previously determined by the Maxam-Gilbert technique⁵. A synthetic oligonucleotide complementary to a variant sequence was protected by mRNA from mung bean nuclease digestion to the same extent as a synthetic oligonucleotide complementary to the pilin signal sequence. Thus, the pilin mRNA which was extended was also the major species expressed by the variants. Furthermore, the synthetic DNA did not prime the synthesis of any complementary DNA in RNA preparations from P⁻ B-series variants.



compensating substitution nearby, and vice versa. For example, a comparison of variants 10A and 10C shows that 10C has lost all charged amino acids at positions 79-84, but compensating charges in 10C can be found approximately 10 residues downstream. At the nucleotide level, these substitutions frequently involve two or three changes, both transitions and transversion, in a codon. The carboxy-terminal portion of the protein is highly variable; in addition to amino-acid substitutions, in-frame insertions and deletions of one to four codons occur. The amount of variability here would explain the low immunological cross-reactivity of pili from different strains, as this part is immunodominant over the conserved portion of the protein^{11,18}.

Two regions which bracket this hypervariable (HV) segment are always conserved; these centre around the cysteines at positions 121 and 151, are 13 and 10 amino acids long, respectively, and may represent epitopes common to all gonococcal pili.

In most variant genes, both the SV and HV sequences are entirely different from those of the progenitor. However, pilin variants of line 3 share the same SV gene segment but different HV ones. Thus it is possible to assemble complete genes by mixing SV and HV gene segments, a mechanism which would considerably increase the antigenic repertoire.

The absence of a good animal model for gonococcal infections has hampered the study of *in vivo* pilin antigenic variation.

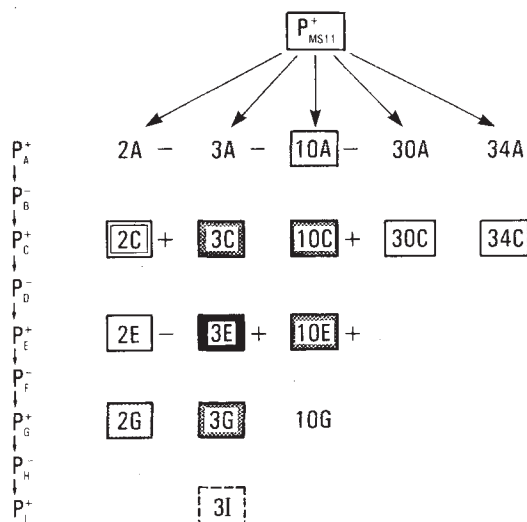


Fig. 2 Left-hand side, a schematic representation of the protocol used for deriving the lines of pilus variants of strain MS11. From the original MS11 strain, which is P⁺, several independent P⁺ colonies were randomly picked. (Piliation can be scored as differences in colony morphology on a plate.) Each P⁺ isolate is the start of a line, each line is given a number and each constitutes the A-series, all of which are identical to the MS11 progenitor. P⁻ revertants are then randomly picked from each P⁺ A-series. These have undergone a P⁺ to P⁻ phase switch once and are now in the B-series. Subsequently, P⁺ backswitchers can be isolated from the different lines in the B-series; they now belong to the C-series. This procedure was used to obtain lines of variants which have undergone the phase switch several times. All manipulations were done in a PII⁻, or opacity⁻, background. Piliation phenotype is given as P⁺ or P⁻, for piliated or non-piliated variants, respectively. The subscripts A-I denote the position within the series to which each variant belongs. The right-hand part of the figure is a schematic representation of the pilus sequence information obtained from the P⁺ variants from lines 2, 3, 10, 30 and 34. Each variant sequence is denoted by the style of the box surrounding the strain number. For example, the progenitor P⁺ MS11 (at the top of the figure), 10A, 30C, 34C and 2E all express the same pilus gene variant. 2C, 3E, 2G and 3I each express a unique gene variant. Those strains which produce pilin antigenically different from that of the progenitor are indicated by + next to the boxes, while those strains which produce pilin identical to that of the progenitor are marked by -. The absence of + or - indicates that the immunological identity of the pilin from these variants has not been determined. Immunological identity is determined by immunodiffusion²⁴ using crude pili preparations and antiserum directed against pili from the MS11 progenitor.

variation seen *in vitro* accurately represent those which control variation *in vivo*.

The P⁺ to P⁻ phase switch involves deletion of pilin sequences from one of the two expression loci⁵. In P⁻ variants 4.1B, 5.1B and 6.1B, we have observed a deletion from only one locus⁵. These variants are non-piliated, and their remaining intact expression site is not transcribed (data not shown). However, *Escherichia coli* harbouring clones of these intact expression sites make pilin transcripts, suggesting that phase variation also involves a switch-induced regulator. Examination of the pilin transcripts derived from *E. coli* shows that sequence variation in 4.1B, 5.1B and 6.1B has occurred during the P⁺ to P⁻ switch (Fig. 1b). The new gene variants have all the characteristics observed for the other P⁺ backswitchers and, therefore, probably encode functional pilins. These observations suggest that pilus phase and antigenic variation are closely, if not obligately, linked.

Our data strongly suggest that the pilus of *N. gonorrhoeae* can undergo extensive antigenic variation. The mixing of SV with HV gene segments predicts that the silent loci will contain truncated pilin sequences. Thus, antigenic variation would involve a multiple-step recombination event in which sections of the pilin gene are joined successfully. The generation of new pilin variants may involve those mechanisms postulated for the generation of immunoglobulin diversity (see refs 20, 21), that is, somatic mutations in the silent copies, unequal crossing-over and/or imprecise joining of SV and HV gene segments (see Fig. 1a, line 3). Also, the tendency of MS11 derivatives to express certain variant sequences could be explained by a mechanism similar to that governing the preferential use of certain joining sequences in pre-B-cell immunoglobulin heavy-chain switching²². Pilin sequences close to those involved in the initial pairing event between the silent and expression loci would be expressed preferentially.

Whether other bacterial pili undergo extensive antigenic variation is unknown. The *Moraxella bovis* chromosome contains multiple pilin gene sequences (C. Marrs, personal communication). *E. coli* K-88 pilus²³ and P-specific pilus (S. Normark, personal communication) variation involves in-frame insertions and deletions as well as single-base changes in specific regions of the structural gene. Thus, the mechanisms used by bacteria to generate immunologically variant surface proteins may be very similar, if not identical.

We thank Emily Chen for making the oligonucleotides and Beth Larsen for typing the manuscript. This work was supported by NIH grant A120845, NSF grant PCM8340588 and MFR grant Dnr 5428.

Received 11 December 1984; accepted 22 February 1985.

- Swanson, J. *J. exp. Med.* **127**, 571-589 (1973).
- Ward, M. E., Watt, P. J. & Robertson, J. N. *J. infect. Dis.* **129**, 650-659 (1974).
- James-Holmquest, A. N., Swanson, J., Buchanan, T. M., Wende, R. D. & Williams, R. P. *Infect. Immun.* **9**, 897-902 (1975).
- Buchanan, T. M. & Pierce, W. A. *Infect. Immun.* **13**, 1483-1489 (1976).
- Myers, T. F., Billyard, E., Haas, R., Storzbach, S. & So, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6110-6114 (1984).
- Herskowitz, I. *Curr. Topics dev. Biol.* **18**, 1-14 (1983).
- Buchanan, T. M. *J. exp. Med.* **141**, 1470-1475 (1975).
- Robertson, J. N., Vincent, P. & Ward, M. E. *J. gen. Microbiol.* **102**, 162-177 (1977).
- Brinton, C. C. *et al.* in *Immunobiology of Neisseria gonorrhoeae* (eds Brooks, G., Gotschlich, E. C., Holmes, K. K., Sawyer, S. & Young, F.) 155-178 (American Society for Microbiology, Washington, DC, 1978).
- Lambden, P. R. *J. gen. Microbiol.* **128**, 2105-2111 (1982).
- Virji, M. & Heckels, J. E. *J. gen. Microbiol.* **129**, 2761-2768 (1983).
- Rothbard, J. B., Fernandez, R. & Schoolnik, G. K. *J. exp. Med.* **160**, 208-221 (1984).
- Hermosdon, M. A., Chen, K. C. S. & Buchanan, T. M. *Biochemistry* **17**, 442-445 (1978).
- Sastry, P. A., Pearlstone, J. R., Smillie, L. B. & Paranchych, W. *FEBS Lett.* **151**, 253-256 (1983).
- McKern, N. M., O'Donnell, I. J., Inglis, A. S., Stewart, D. J. & Clark, B. L. *FEBS Lett.* **164**, 149-153 (1983).
- Hamlyn, P. H., Gait, M. J. & Milstein, C. *Nucleic Acids Res.* **9**, 4485-4492 (1981).
- Segal, E., Billyard, E., So, M., Storzbach, S. & Meyer, T. F. *Cell* **41** (in the press).
- Schoolnik, G. K., Fernandez, R., Tai, J. Y., Rothbard, J. & Gotschlich, E. C. *J. exp. Med.* **159**, 1351-1370 (1984).
- Fletcher, J. D., Stratton, J., Chandler, C. S. & Sparling, P. F. *CDC Morbid. Mortal. Wkly Rep.* **No. 32**, 273-275 (1983).
- Tonegawa, S. *Nature* **302**, 575-581 (1983).
- Honjo, T. A. *Rev. Immun.* **1**, 499-528 (1983).
- Yancopoulos, G. D. *et al.* *Nature* **311**, 727-733 (1984).
- Josephsen, J., Hansen, F., de Graaf, F. K. & Gaastra, W. *FEBS Lett.* (in the press).
- Weir, D. M. (ed.) *Handbook of Experimental Immunology* (Blackwell Scientific, Oxford, 1979).

Recently, an epidemic of gonorrhoea occurred¹⁹ whose origin could be traced to a single source and whose isolates were shown to be members of the same strain (P. F. Sparling, manuscript in preparation). We have chosen isolates from three patients for our sequence studies. One series (ML1, 2, 3) was obtained from a patient who was repeatedly infected between successful courses of antibiotic therapy. Two other isolates, LP1 and WG3, were from different patients. Our data (Fig. 1c) show that pilin antigenic variation did occur *in vivo* during the epidemic. It also occurred in the ML isolates obtained from a single individual at various times. However, because antibiotic therapy was instigated between each bout of infection and no sera were available from this patient, we cannot determine whether pilus antigenic variation in these cases was caused by the selective pressure exerted by the host immune response or whether these individuals were infected by partners carrying an already variant strain. However, it is clear that a single gonococcal source expressing one pilus type can, during the course of the epidemic, give rise to isolates which now produce variant pili. The C, SV and HV regions defined by the MS11 sequences are also observed in these variants. Thus, the events contributing to the antigenic

The nonlinear Poisson-Boltzmann equation

SPITZER¹ has reminded us that the nonlinear Poisson-Boltzmann equation (NPBE) contains a logical inconsistency. An early incisive analysis of this inconsistency was given by Onsager² 50 years ago. Unhappily, more recently, a considerable confusion^{3,4} has accrued concerning the nature and consequences of the inconsistency. Spitzer's¹ statement that "the NPBE violates the important electrostatic theorem of additivity of electrostatic potentials" is simply incorrect, and his recommendation that "the linearization of the NPBE is desirable, in principle" is entirely inappropriate.

First we consider the nature of the inconsistency. Onsager² was concerned with the microscopic NPBE as it is used to describe the potential generated by a single ion in solution that is screened by the response to it of all the other ions in the electrolyte. As Onsager makes clear, the inconsistency derives from the replacement of the potential of mean force, $w_{ji}(r)$, in the correct Boltzmann relation (his equation (21); ref. 2), by the simpler electrostatic potential energy $e_i\psi_j(r)$ to obtain what is finally an inconsistent Boltzmann relation (his equation (13); ref. 2). The replacement is made in order to write a one-particle NPBE. Note that neither the Poisson equation nor the superposition theorem of electrostatics as such is germane to the inconsistency studied by Onsager. The confusion is compounded by a supposed connection of the linear superposition of electrostatic potentials to a linear dependence of space charge ρ on potential ψ (refs 1,3). In actuality, the linear superposition theorem of electrostatics⁴ holds for an arbitrary distribution of space charge $\rho(r)$. Thus, in the analogue to the NPBE appropriate to Fermi-Dirac statistics, namely, the Thomas-Fermi equation⁵, $\rho \propto \psi^{3/2}$. For space-charge-limited currents in a planar vacuum diode⁶, $\rho \propto \psi^{-1/2}$; in the analogous solid-state diode⁷, $\rho \propto \psi^{-1/3}$, and so on.

We next take up the consequences of the inconsistency. As noted above, the inconsistency refers directly to the microscopic NPBE. However, it is important to note that most of the literature on the NPBE since World War II, including the Spitzer letter¹, is not concerned with this equation but rather with the macroscopic NPBE. This latter equation describes the spatial dependence of the average potential generated by relatively large, charged surface(s) in contact with the electrolyte. This local average potential is what remains after all microscopic fluctuations have been averaged out. For the macroscopic NPBE a new phenomenon comes into play when linearization finally fails (that is, at high surface-charge concentrations), which is the formation of a counterion condensate close to the surface⁸.

Fixman⁹, in a theoretical study of the problem within the framework of statistical mechanics, pointed out that substantial errors deep in the condensate lead to far smaller errors outside the condensate. He aptly described this phenomenon as "the buffering action of 'the condensed phase'" (ref. 9). This very pronounced decoupling of a large surface charge from the 'distant' bulk electrolyte is also a direct mathematical consequence of the NPBE, as shown for planar geometry by Verwey and Overbeek¹⁰. The buffering action of the condensate, as predicted by the NPBE, is even more striking for cylindrical and spherical geometries, as shown by numerical studies of this equation by Lampert and Martinelli (to be published elsewhere). Fixman's numerical calculations led him to conclude that the NPBE is acceptable⁹ up to local, condensate concentrations as high as 2–3 M, provided that the bulk salt concentration does not exceed ~ 0.1 M, which happens to be, approximately, that for physiological saline. It is also relevant here and perhaps not widely known that some 15 yr after his incisive criticism of the microscopic NPBE, Onsager used the macroscopic NPBE to calculate 'The effects of shape on the interaction of colloidal particles'¹¹.

Spitzer¹ alludes to the difficulty of testing experimentally the Gouy-Chapman-Stern (GCS) model for incorporating adsorption into double-layer theory, as constructed with the NPBE. If that is the situation in electrochemistry and colloid chemistry, then matters are improved considerably in biophysical studies of charged phospholipid bilayer membranes immersed in an aqueous solution. McLaughlin *et al.*¹² showed, by adding small concentrations of divalent cations to a host solution of monovalent ions, that the observed shift in membrane potential disagreed strongly with predictions of linearized PB theory, but agreed adequately with GCS theory. In a recent critique of the work of several groups using various experimental techniques, McLaughlin concludes¹³ "that fortunately, the GCS theory appears to be a good first approximation for describing the adsorption of cations to bilayer membranes".

M.L. is on medical leave from Princeton University.

MURRAY LAMPERT

Department of Electrical Engineering
and Computer Science,
Princeton University,
Princeton, New Jersey 08544, USA

1. Spitzer, J. J. *Nature* **310**, 396–397 (1984).
2. Onsager, L. *Chem. Rev.* **13**, 73–89 (1933).
3. Bockris, J. O'M. & Reddy, A. K. N. *Modern Electrochemistry*, 247 (Plenum, New York, 1970).
4. Jackson, J. D. *Classical Electrodynamics*, 13–14 (Wiley, New York, 1962).
5. Messiah, A. *Quantum Mechanics*, 616 (North-Holland, Amsterdam, 1966).
6. Smythe, W. R. *Static and Dynamic Electricity* 3rd edn, 272–274 (McGraw-Hill, New York, 1968).

©1985 Nature Publishing Group

7. Lampert, M. A. & Mark, P. *Current Injection in Solids*, 45 (Academic, New York, 1970).
8. Manning, G. S. *J. chem. Phys.* **51**, 924–933 (1969).
9. Fixman, M. *J. chem. Phys.* **70**, 4995–5005 (1979).
10. Verwey, E. J. W. & Overbeek, J. Th. G. *Theory of the Stability of Lyophobic Colloids* (Elsevier, Amsterdam, 1948).
11. Onsager, L. *Ann. N. Y. Acad. Sci.* **51**, 627–659 (1949).
12. McLaughlin, S. G. A., Szabo, G. & Eisenman, G. *J. gen. Physiol.* **58**, 667–687 (1971).
13. McLaughlin, S. in *Membranes and Transport* (ed. Martonosi, N.) Ch. 6 (Plenum, New York, 1982).

Homologies in the avian tarsus

MCGOWAN has recently presented evidence from the avian tarsal joint for a wide taxonomic separation between ratite and carinate birds. He implies that the pretibial bone of carinates is formed by ossification of a portion of the cartilaginous precursor of the calcaneum (a "calcaneal spur"), while in ratites the comparable structure is an ascending process of the astragalus. However, the youngest specimen he described was a 12-day-old quail embryo. Examination of a large series of cleared and stained chicken embryos shows that the 7-day-old embryos have separate cartilaginous calcania, astragali and tibiae, but no ascending processes or pretibial bones. Eight-day-old chicken embryos show a separate triangular pretibial chondrification above the calcaneum which does not appear to be part of either the astragalus or the calcaneum. On day 9 the pretibial, astragalar and calcaneal cartilages fuse. The pretibial cartilage appears to fuse with both the calcaneal and astragalar cartilages. Only in the fused state could one mistake the pretibial chondrification for an extension of the calcaneum. By day 17 the pretibial bone (calcaneal spur of McGowan¹) has begun to ossify; at this point it is the only tarsal bone undergoing ossification (Fig. 2b–d of ref. 1). Thus, the pretibial bone is the last carinate tarsal element to chondrify and the first to ossify.

McGowan¹ shows that the first bone to ossify in the tarsal complex of ratites is the "ascending process". We believe that this concordance between the ossification of the ascending process of ratites and the pretibial bone of carinates would be hard to explain if they are not the same bone. This is especially true when we consider that the calcaneum is the last bone to ossify in McGowan's preparations.

McGowan¹ describes the "ascending process" of ratites as an ossification with a disk-like base and separate from the astragalus. He describes an early stage in the development of the distal end of the tibiotarsus as comprising a medial ossification and a lateral cartilaginous area. The lateral cartilaginous area has a dorsal spur which he considers to be "homologous with the precursor of the carinate pretibial bone". The enlargement of the astragalus found in some ratites might be a derived character state related to neoteny. The earlier ossification of the astragalar portion of the joint may increase the relative

importance of the astragalus in highly neotenic birds. Despite McGowan's statements, the pretibial bone seems to represent both a separate chondrification and ossification in carinates; we believe that this is also the case in ratites and that the difference (if it exists) between ratites and carinates is the result of varying amounts of overlap with the calcaneum and the astragalus. Because of the extreme lateral placement of the pretibial bone in some carinates, it may be confined to the region above the calcaneum (see Fig. 11 of Martin *et al.*²), and after the chondrification has fused with the calcaneum it may resemble a 'spur' of that bone. On the other hand, falcons have the pretibial bone very centrally situated as in the ostrich, and most, if not all, of the fusion of the distal end of the pretibial bone is with the astragalus (Fig. 179 of ref. 3). A juvenile swan, *Cygnus buccinator* [Kansas University 79260], shows a pretibial bone ossifying separately from the calcaneum and partially overlapping the astragalus; except for its slightly more lateral position, it corresponds to McGowan's description of the "astragalar process of the ostrich" which is a "well-ossified ascending process, expanding distally into a disk of bone".¹ We believe that the present evidence indicates that ratites and carinates have a homologous ossified ascending process (pretibial bone) separate from both calcaneum and astragalus, but that the position of this process may vary. We believe that, as originally suggested, the pretibial bone is a neomorph ossification providing a derived character for birds. The late chondrification of this element supports such an interpretation. The Mesozoic birds (*Archaeopteryx*, *Enaliornis*, *Baptornis* and *Hesperornis*) have the pretibial bone laterally placed and fused to the calcaneum as in carinates, but not as in theropods (Fig. 1b of ref. 1), and we would therefore caution against arguing that the pretibial bone of carinates is a derived character state within the Class Aves.

To view these results as in any way confirming the theropod origin of birds would be to beg the question. The only way McGowan could conclude that the ratite tarsal condition is more primitive than that of the carinates is by already assuming that theropod saurischians are the out-group most proximate to birds; this approach does not address the fact that pseudosuchians and crocodilians are also contested nearest out-groups⁴⁻⁷. McGowan's own interpretation of his material would indicate that the bulk of the known birds (the carinates) do not have a theropod-like ascending process.

We thank C. Wyttenbach and J. Garrison for the use of cleared and stained chicken embryos. H.-P. Schultze and M. Mengel critically read the manuscript, and R. Hermann typed it. Support for work on bird origins has been funded by NSF DEB 7821432, KU 3251-5038, and National Geographic grant 2228-80.

L. D. MARTIN
J. D. STEWART

*Journal of Natural History and
Department of Systematics
and Ecology,
University of Kansas,
Lawrence, Kansas 66045, USA*

1. McGowan, C. *Nature* **307**, 733-735 (1984).
2. Martin, L. D., Stewart, J. D. & Whetstone, K. N. *Auk* **97**, 86-93 (1980).
3. Jollie, M. *Evol. Theory* **3**, 1-141 (1977).
4. Benton, M. J. *Nature* **305**, 99-100 (1983).
5. Martin, L. D. *Current Ornithology*, 105-129 (1983).
6. Whetstone, K. N. & Whybrow, P. J. *Occ. Pap. Mus. nat. Hist. Univ. Kansas* **106**, 1-37 (1983).
7. Tarsitazno, S. & Hecht, M. K. *Zool. J. Linn. Soc.* **69**, 149-189 (1980).

MCGOWAN REPLIES—According to Martin and Stewart, the ascending process of the carinate tarsus commences as a chondrification independent of the cartilaginous astragalus and calcaneum. This subsequently ossifies as a separate element, and a similar ontogeny is believed to occur in ratites. The pretibial bone thus formed, a new element which is not homologous with the ascending process of theropods, provides a shared derived character for birds.

My X rays of juvenile ostriches (ref. 1, Fig. 3b-c) confirmed Huxley's² belief that ratites, like theropods, have an ascending process that is continuous with the bony astragalus. Whether this process develops in ratites as an integral part of the astragalus, or, as Martin and Stewart contend, as an independent chondrification, and ossification, can only be determined embryonically. The earliest ratite embryos available to me were close to hatching, at which stage there is a prominent bony ascending process expanded distally into a disk of bone embedded in cartilage. That this entire structure is the astragalus, and not an independent ossification, is confirmed by examining early ostrich and emu embryos³; these show that the cartilaginous astragalus is drawn up into an ascending process, as Parker⁴ described for the kiwi, and this subsequently ossifies to give the bony ascending process. Ossification commences at the tip of the cartilaginous process, spreads distally and eventually extends across the entire width of the cartilaginous astragalus, as shown in my X rays of juveniles. Ossification of the corresponding process in carinates is confined to the process itself (ref. 1, fig. 2b-e).

I have examined swan embryos (*Cygnus olor*), and find that the pretibial bone develops as an ossification of a cartilaginous spur, as described for other carinates¹. Inasmuch as the pretibial bone ossifies separately from the bony calcaneum, my observations¹ do not conflict with those of Martin and Stewart, and I believe that if they examined later stages of swan development they would find that the pretibial bone eventually fused with the calcaneum, as it does in the chicken (ref. 1 fig. 2f).

One area of common ground I now share with Martin and Stewart pertains to

the homology of the ascending process in ratites and carinates, but our conclusions differ radically. An examination of early duck embryos (days 8-11 of incubation) revealed a dorsal cartilaginous process from the lateral aspect of the astragalus, similar to that seen in early emu embryos (15-20 days)³. This early association with the astragalus is soon obscured by the fusion of the cartilaginous astragalus and calcaneum, and from then on the process develops in close association with the calcaneum, as described previously¹. Similar observations for gulls lead me to believe that the carinate pretibial bone is an astragalar derivative, supporting its homology with the ascending process of ratites.

Ratites, like theropods, have a bony ascending process that is continuous with the astragalus, but carinates have a derived condition where the process fuses with the bony calcaneum. Martin and Stewart deny that the (predominantly lateral) carinate pretibial bone is a derived character within birds, noting that the ascending process is laterally placed in the earliest bird, *Archaeopteryx*, but this is the case in theropods¹; the significant point is not its orientation but whether it is continuous with the astragalus or the calcaneum.

My conclusion that the ratite tarsus is more primitive than that of carinates does, of course, hinge upon theropods being phylogenetically closest to birds, a position I took at the outset with my reference to Ostrom's work⁵. Citing a review by Benton⁶, Martin and Stewart point out that pseudosuchians and crocodilians have also been proposed as the nearest avian out-groups. Their own bias for the crocodilian hypothesis is contained in a paper by Martin *et al.*⁷. Note, however, that Walker⁸, the leading proponent of the crocodilian hypothesis, has recently reassessed the evidence, concluding that the hypothesis for a common ancestry of crocodiles and birds (above the thecodont level) should be rejected.

I thank H. Lumsden for supplying swan embryos; Dr A. D. Walker for access to his unpublished work; A. J. Baker for reading the manuscript; and the Natural Sciences and Engineering Research Council of Canada for funding support.

C. MCGOWAN

*Department of Vertebrate Palaeontology,
Royal Ontario Museum, Toronto
and Department of Zoology,
University of Toronto,
Toronto, Ontario,
Canada M5S 1A1*

1. McGowan, C. *Nature* **307**, 733-735 (1984).
2. Huxley, T. H. *Q. J. geol. Soc. Lond.* **26**, 12-31 (1870).
3. McGowan, C. *J. Zool.* (in press).
4. Parker, T. J. *Phil. Trans. R. Soc. Lond.* **182**, 25-134 (1892).
5. Ostrom, J. H. *Nature* **242**, 136 (1973); *Q. Rev. Biol.* **49**, 27-47 (1974); *A. Rev. Earth planet. Sci.* **3**, 55-57 (1975); *Proc. Cent. natn. Rech. Sci. Colloq. internat.* **218**, 519-532 (1975); *Biol. J. Linn. Soc. Lond.* **8**, 91-182 (1976).
6. Benton, M. J. *Nature* **305**, 99-100 (1983).
7. Martin, L. D., Stewart, J. D. & Whetstone, K. N. *Auk* **97**, 86-93 (1980).
8. Walker, A. D. *Nature* **237**, 257-263 (1972); in *Problems in Vertebrate Evolution* (eds Andrews, S. M., Miles, R. S. & Walker, A. D.) 319-358 (Academic, London, 1977).

In the aftermath

P.M. Kelly

Nuclear Winter: The Human and Environmental Consequences of Nuclear War.

By Mark A. Harwell.

Springer-Verlag: 1984. Pp.179. DM 54.

WITH a fresh spate of books on the "nuclear winter" about to reach the marketplace, it is appropriate that Mark Harwell's sober and detailed account of the biological aspects of the subject should be available first. It deals not only with the effects of a possible nuclear winter but also with the more direct, and inevitable, blast, fire and radiation damage caused by nuclear explosions. The book is based on a technical support document produced, in collaboration with other biologists and ecologists, for the conference "The World After Nuclear War".* It was at this meeting, held in Washington DC in late 1983 and the culmination of a series of exercises designed to assess the scientific validity of the nuclear winter theory, that the first detailed findings concerning the possibility of widespread cooling and concomitant damage to the biosphere were announced.

Shortly after the Washington conference, two summary papers were published in *Science* (222, 1,283-1,300). One dealt with the atmospheric component of the theory; the other with the biological aspects. To those aware of developments only through these papers or through brief, sometimes opinionated, accounts in the scientific press and the news media, there seemed to be a lot of fuss about a rather speculative notion. This was a mistaken impression. A great deal of carefully documented background research was available to anyone who cared to look for it, as Harwell's account of the rigorous analysis supporting the biologists' case clearly demonstrates.

But how certain is the atmospheric scientists' case? Since 1983, the principal conclusion of the initial investigators, the possibility that a nuclear war would cause a major climatic excursion — an abrupt, marked but short-lived change — has been confirmed by other researchers and by two review projects mounted by the US National Academy of Sciences (NAS) and the Royal Society of Canada (RSC). There are many areas of uncertainty, and the NAS and RSC reports carefully define them, but the risk of a nuclear winter now seems undeniable.

The NAS committee concentrated on the atmospheric aspects of the problem, but noted that the impact of the climatic excursion

on those who survived the initial effects of a nuclear exchange and on the biosphere as a whole "deserves careful independent study". Progress on the physical side has been rapid since the Washington conference. Many points in the original investigations have been clarified and the computer simulations of the nuclear winter effect have been substantially improved. On the biological side, progress has been slower. In part, this has been an unfortunate by-product of the understandable reluctance of atmospheric scientists to provide the type of detailed information on climate needed for rigorous impact assessment. Computer models of the climatic system tend not to be sufficiently accurate on the spatial and time scales of concern.

One solution to this problem is to use scenarios for possible, in the sense of physically plausible, nuclear winter climates in much the same way as atmospheric scientists assessing the climatic effects use a range of strategically plausible nuclear war scenarios. This is Harwell's philosophy. While concentrating on the consequences of a nuclear exchange in the 5,000 megaton range, the book provides the data necessary to assess the implications of any scale of exchange (or nuclear winter).

Beginning with the development of a nuclear war scenario, which provides the input data for the "consequence analyses", the direct effects of blast, fire and fallout (both local and global) are considered first. This area has been addressed in depth in a number of studies, and Harwell's main contribution is to quantify the impact in terms of casualties and damage to the biosphere in the United States. This section is an excellent introduction to the literature on these aspects of the post-nuclear world. The concept of the nuclear winter is then introduced and a detailed discussion of its effects on the biosphere forms the main part of the rest of the book.

In considering the consequences of the principal assaults — temperature and sunlight reductions, radioactive fallout, and the loss of technical and social support for agriculture — Harwell assesses the net impact on the human population and on the flora and fauna, the food supply for the survivors. Using numerical models whenever available, the impact is quantified in terms of human casualties and food availability. The emphasis is, for the most part, on the northern middle latitude combat zone, and grassland, forest, freshwater and estuarine ecosystems are subject to particular scrutiny.

Inevitably, some conclusions are more speculative than others, but the reader is given all the information necessary for a personal assessment of the uncertainties.

In the midst of this description of the gross distortion and near-destruction of the ecology of the combat zone, some surprising insights into possible diets emerge. Ducks, for example, through metabolic generation of heat, could survive temperatures of -40°C for a couple of weeks (in the unlikely event that enough food was available). The analysis ends with a brief section entitled "Recovery Processes". I find the use of the word "recovery" in this context somewhat inappropriate. "Coping" and "adaptation" seem more suitable terms given that many aspects of the pre-nuclear environment (and civilization) might never be recovered.

Harwell also discusses possible psychological responses to the devastation that would occur, but wisely steers clear of dogmatic statements. Similarly, the nature of the fragmented post-nuclear social structure is covered but no firm conclusions are reached. These areas have yet to be considered in any depth but, as the author observes, such factors might well be crucial in determining whether or not use could be made of what limited food supplies would be available. Of particular concern, on the individual level, is the prospect of "psychic numbing" through which unacceptable aspects of reality are shut out and which could severely handicap the ability to survive in the aftermath of nuclear war. Would any form of social organization, whether voluntary or enforced, be possible in such circumstances?

The final section, "Summary of Consequences", is a sobering account of the conclusions of the study: the state of the world after nuclear war. It left one question uppermost in my mind. What can be done to provide viable civil defence and support for those who survived the initial attacks in the face of the multiple hazards of the post-nuclear world?

Harwell's stated intention is "to provide a more comprehensive view of what the world would be like to the immediate survivors of a nuclear war" (p.xv). With the assistance of some (rather harrowing) imaginative effort on the part of the reader, a portrait of the world after nuclear war does indeed emerge from the text. But the main strength of the book lies in the clear, dispassionate accumulation of the technical information needed to assess the implications of the use of nuclear weapons. Harwell has produced an exhaustive, wide-ranging and, in many respects, pioneering treatise which will be an invaluable reference tool for anyone concerned about nuclear issues. □

*The book resulting from the conference, *The Cold and the Dark*, has been published in paperback in Britain by Sidgwick & Jackson under the title *The Nuclear Winter: The World After Nuclear War* (price £5.95).

Fundamentals of the fourth state

C.N. Lashmore-Davies

Handbook of Plasma Physics, Vols I and II.

Edited by A.A. Galeev and R.N. Sudan.
North-Holland: 1983/1984. Vol. I pp.751, \$180.75, Dfl.425. Vol. II pp.850, \$190.50, Dfl. 495.

AT A TIME when the growth in the publication of papers on plasma physics shows no sign of slowing, the appearance of two volumes which set out the basic ideas of the subject and provide summaries of all the main lines of research is particularly welcome. These two books are worthy successors to the Leontovich reviews and the Simon and Thompson series, and have the advantage of being published together. They contain 33 articles, 17 of which were contributed by 19 authors from the United States, 15 by 12 authors from the Soviet Union and one by two Japanese authors.

Volume I deals with the principles of plasma physics, and is divided into four parts. The first covers the motion of charged particles in electromagnetic fields, atomic collision processes relevant to plasmas and the various radiation phenomena which can occur in plasmas, concluding with an illuminating discussion of the different modes of description of plasmas for low frequency phenomena and an account of the subtleties of collisional transport in magnetically confined plasmas.

The remaining parts of Vol. I deal with the collective nature of plasmas. An important feature of a plasma is the great variety of waves it can support. The nature of the linearized wave motions allowed by ideal MHD and the effect of allowing for a distribution of particle velocities are both described. Also included are treatments of fluctuations, geometrical optics and the conversion of one linear wave into another in an inhomogeneous plasma.

When there is a reservoir of free energy in a plasma some of these wave modes can become unstable resulting in a great variety of instabilities. The many aspects of plasma instabilities from ideal and resistive MHD, the kinetic instabilities driven by the pressure gradient of a confined plasma, non-thermal particle distributions in velocity space and whether a given instability behaves like an amplifier or an oscillator are all systematically considered. An inevitable consequence of instability is that a plasma moves into a state in which its motion is non-linear. Since a plasma can support so many different wave motions it has been conjectured that a state of weak turbulence might exist, and in the final part of Vol. I the concepts behind this theory are lucidly set forth.

Volume II follows on from its companion, containing Parts 5-9, and is mainly

concerned with non-linear consequences. In many practical situations an enhanced level of long wavelength Langmuir waves can be generated (for example by an electron beam or a laser) which, according to weak turbulence theory, should relax to a coherent Langmuir oscillation. The resolution of this apparent paradox and the existence of three-dimensional, localized, high-frequency electric fields are beautifully described by the discoverer of the phenomenon. Also discussed are the spectra resulting from weak turbulence and the numerical modelling of strong Langmuir turbulence.

In hot plasmas the effect of binary collisions is often insufficient to explain various transport processes. The simplest of these phenomena is the dissipation of a current in an unmagnetized plasma, due to the generation of waves rather than inter-particle collisions, although even this apparently simple example is not fully understood. Other examples considered are the dissipation of intense laser light and finite amplitude waves in plasmas through the non-linear generation of waves. Similar but far more complicated situations occur in the cross-field diffusion of particles and energy in magnetically confined plasmas. This complex issue is surveyed, together with an attempt at a comprehensive statistical description of plasma turbulence. These effects are generally referred to as anomalous although in many applications they represent the normal situation rather than the exception.

Part 8 consists of two excellent reviews on the diagnostic methods upon which the subject of plasma physics rests; the first is concerned with radiative techniques and the second with the methods used in hot magnetically confined plasmas, complementing those discussed in the first article. Volume II also contains an informative account of the methods of particle simulation by computer and a thorough discussion of the equilibrium and stability of non-neutral relativistic electron beams and their interaction with plasma.

Together, the books satisfy the aims of the general editors in encompassing most of the present-day knowledge of plasma physics, albeit from a user point of view. They are not suitable for someone requiring an introduction to the subject; although they take the reader to the frontiers of current research, they do not cover the general topic of the foundations of plasma kinetic theory nor do they contain more than a very brief account of the fundamental and fascinating subject of Landau damping. But those actively involved in research on plasma physics will find these volumes invaluable, as will experienced workers in other fields wishing to acquaint themselves with a particular aspect of the subject. □

C.N. Lashmore-Davies is at the Culham Laboratory, Abingdon, Oxfordshire (Euratom/UKAEA Fusion Association).

©1985 Nature Publishing Group

Off the record

John Parkington

From Hunters to Farmers: The Causes and Consequences of Food Production in Africa.

Edited by J. Desmond Clark and Steven A. Brandt.

University of California Press: 1984.
Pp.433. \$63.25, £52.25.

AFRICA, as one of my mentors once sagely remarked, is a very big place. It encompasses environments as diverse as the sand seas of the Sahara and Namib, the thirlands of the Kalahari and Sahel, the heath communities of the Cape and Mahgreb, the savannahs of Serengeti and the tropical forests of Zaire. To understand the social and ecological events leading to and following from the transition from food gathering to food production over such a vast and varied area is a Herculean task. Yet archaeologists are nothing if not optimists, and the goal of *From Hunters to Farmers* is nothing less than this. A collection of 30 papers, it sets out to provide a synthesis of current knowledge and is, in effect, the published proceedings of a symposium held in Los Angeles in 1978.

The volume is divided into four sections, the first consisting of three general papers which present a depressing but honest picture of how little we know about past climates, past population distributions and the circumstances surrounding the adoption of food production in Africa. Most of the contributions fall into the second section and are regional summaries of greater or lesser geographical scope. I wondered at this point what distinguishes papers that are quoted long after publication. Do they contain dazzling theoretical insights or essential primary observations? Or do they reflect the spirit of their times? Regional reconstructions based on modest but accumulating data sets are almost by definition transitory phenomena. We all write and then rewrite them.

To my mind the most interesting section is the third, in which the effects of the acquisition of domestic plants and animals by Kalahari San are documented. These three papers combine empirical observations with theoretical perspectives and will probably have longer lives than the regional accounts. Because they contribute to the realization that the San are, and have been, far from isolated from agriculturalists and pastoralists, they will have immediate im-

New in paperback

● *The Creative Computer: Machine Intelligence and Human Knowledge* by Donald Michie and Rory Johnston. Publisher is Penguin, price is £6.95. For review see *Nature* 313, 77 (1985).

● *When the Snakes Awake: Animals and Earthquake Prediction* by Helmut Tributsch. Publisher is MIT Press, price is \$8.85, £9.25. For review see *Nature* 302, 763 (1983).

pact. The only paper in the fourth section describes the difficulties involved in palaeodemographic reconstruction and provides a good overview of what we would like to know but usually find hard to document archaeologically.

As the editors undoubtedly expected, this book is more valuable in what it shows up rather than what it shows off. The contributors have demonstrated that we know very little of the climatic and environmental circumstances that preceded the appearance of domesticates, that the preservation of reasonable samples of plants and animal bones in archaeological contexts is variable but generally poor, that we have very inadequate dating of critical changes, and that models for understanding *how* and *why* people shift into agriculture or pastoralism are only now emerging. More seriously we see an archaeological record that is so residual and fragmentary that it becomes malleable and an easy victim for a predatory hypothesis. In these circumstances it is extremely tempting to "fit" the archaeological evidence

onto superficially more substantial linguistic histories, climatic reconstructions or, more commonly, ethnographic models.

Several of the contributors (particularly but not only Stahl, Ambrose, Klein, Gifford-Gonzalez, and McIntosh and McIntosh) make the point that archaeological observations have to generate some internal and independent consistency. This comes only from sustained multidisciplinary research into regional sequences of events, something that few parts of Africa have so far enjoyed. What is known is largely in this book. Although I would like to have read more discussion of theory and method, to have seen more concern with social than ecological issues and to have heard more talk of process, I recognize and recommend here a valuable summary of our current understanding of "the causes and consequences of food production in Africa". □

John Parkington is in the Department of Archaeology at the University of Cape Town.

Search for hay in a needlestack

Harry Hertz

Organic Trace Analysis.

By K. Beyermann.

Ellis Horwood/Halsted: 1984. Pp.365. £35, \$74.95.

THE search for trace organic substances in complex inorganic and organic matrices is one of the most challenging problems facing chemists. It pervades all aspects of our lives: the water we drink, the food we eat, the environment in which we work and live, the quality of our manufactured goods and many aspects of our day-to-day health. Klaus Beyermann's *Organic Trace Analysis*, a generally very readable translation of the original German edition, is the first comprehensive description of this developing field of analytical chemistry.

Beyermann has aimed the treatise at a diverse audience, from the professional analyst to "science administrators, biologists, politicians [and] forensic experts". While this is an admirable goal, it has led to occasional identity crises and uneven presentation of various topics. Nevertheless, if readers are willing to pick and choose what to read they will find that the book does present an effective overview of the field.

An outstanding feature is the references, some 3,000 of them. Much of the information on previous studies is reported in tables, permitting easy access to the literature. Furthermore, all the tabular entries (as well as the text itself) are extremely well indexed in the back of the book, making

this an outstanding rapid-reference volume. The order of presentation follows the logical sequence of analysis: from general aspects through sampling, sample handling, component separations and detection.

There are, however, some flaws. The book as a whole is less evaluative of the literature than one would like, the author usually choosing to cite relevant works without any commentary. Section 2.4, which deals mainly with statistics, was not contributed by Beyermann and is not as well written as the rest of the book. The level of detail in this introductory section is perhaps excessive, while the references here are more parochial than they are elsewhere; this could dissuade some readers from continuing. And the topic of sampling and storage is treated in a manner which belies the problem in this area of assuring analyses which accurately represent the native (*in situ*) concentration of the species of interest.

As with any rapidly evolving field, by the time a book has been through the publication mill new discoveries have appeared and contributed to the science being discussed. Here consideration of such techniques as fast atom bombardment in mass spectrometry and the impact of monoclonal antibodies on immunoassay is not included, and thus the book already has a somewhat dated look in parts. For all that, the publication of *Organic Trace Analysis* is undoubtedly a milestone in analytical chemistry. It is a handy reference volume for the practising analyst and a valuable introduction for the generalist who is prepared selectively to read sections of the book. □

Harry Hertz is Director of the Center for Analytical Chemistry at the US National Bureau of Standards, Gaithersburg, Maryland.

©1985 Nature Publishing Group

Angles on aphids

M. Llewellyn

Aphid Ecology.

By A.F.G. Dixon.

Blackie/Methuen: 1985. Pp.157. £17.95, \$39.95.

APHIDS are specialist phloem-feeding insects that have received world-wide attention, both because of their importance as agricultural pests and because of the intricacy and diversity of their biology. A book on aphid ecology by a leading research worker in both fields will thus attract attention from a number of angles. Satisfying all potential readers — especially in a mere 136 pages of text — is a tall order, but Professor Dixon has succeeded remarkably well.

This book is as much about the biochemistry, physiology, genetics and evolution of aphids as about patterns and processes of ecology; it contains not only information about a specialist taxon of insects but constitutes an integrated approach to ecology which, although often talked about, is rarely attempted. Many areas are excellently reviewed. The accounts of such topics as "galling", polymorphism, genetic change, sexual morph production and host-alternation are a delight, and a large literature on the population dynamics of several species has been crystallized into a fascinating and informative chapter. A few subjects do not catch the imagination, however; the section on host plant finding and nutrition, for example, is not particularly convincing. The book is peppered with the author's own ideas and theories — on aphid size, ovariole number, reproductive strategies and the aphid fauna of the tropics — which may not find universal approval but which will stimulate the work of aphid biologists for some time to come.

There are well over 100 graphical figures, as befits an experimental, quantitative approach to ecology, but I often found them of little value because of the lack of adequate background information. The condensed treatment of some topics, in particular reproductive strategies and aspects of migration, may baffle the general reader, and it is a pity that the space taken by some figures was not used to extend consideration of difficult concepts such as these.

After reading this book few will disagree that a broad approach to ecology, drawing upon diverse areas of biology, is anything other than desirable. Here it has resulted in a challenging text which surely heralds the demise of the specialist and blinkered approach to ecology so prevalent over the past 30 years. □

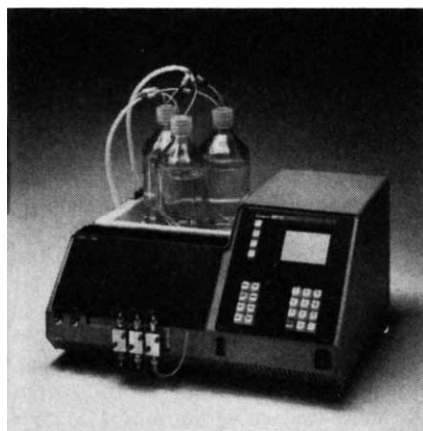
M. Llewellyn is a Lecturer in the Department of Biology at Queen Elizabeth College, University of London.

Forensics and receptor binding

Forensic drug analysis is covered in a new data sheet and a new system simplifies receptor binding assays.

● Latest option for EG&G Ortec's System 5000 energy dispersive spectrometer is the Model 5165 system for digital beam control and X-ray colour mapping. The Model 5165 allows the System 5000 to store and display up to 12 elemental maps of data acquired from an energy dispersive X-ray spectrometer connected to an electron microscope. All 12 maps plus a histogram and an expanded display of one of the maps may be displayed simultaneously. *Reader Service No. 100.*

● Two new polymeric ion-exchange columns have been developed by Interaction Chemicals Inc. for the separation of peptides that are difficult or impossible to separate using silica-based columns. Column efficiencies approach 100,000 plates per metre. The BIO-10 column contains a low capacity material particularly suited for separation of acidic peptides. The BIO-20 column contains a low capacity polymer that will easily separate basic peptides using low ionic strength eluents at elevated pH values. Acidic peptides are not retained on this column, so group separations are readily accomplished. Both polymers are stable from pH 0 to 15 and can be used with most aqueous eluents. *Reader Service No. 101.*



The Model 9600 from Eldex.

● The Eldex Model 9600 is an integrated, programmable solvent delivery system designed for applications ranging from micro-bore HPLC separations to high-resolution preparative scale purification. The Model 9600 is designed to generate ternary gradients with pulse-free solvent delivery and is fitted with ports for interfacing with other chromatography peripherals. *Reader Service No. 102.*

These notes are based on information provided by the manufacturers. For further details circle the appropriate numbers on the Reader Service Card bound inside the journal.

GOW-MAC INSTRUMENT CO. PEAK TO PEAK APPLIED SEPARATION TECHNIQUES

FIELD OF INTEREST: Rapid Identification of Anesthetic Gases

GOW-MAC INSTRUMENT CO.

PEAK TO PEAK APPLIED SEPARATION TECHNIQUES

APPLICATION:

Gas Chromatographic Analysis of Aliphatic and Aromatic Hydrocarbons in Admixture

FIELD OF INTEREST:

Analysis of solvent emissions in: gas phase analysis (e.g. auto exhaust analysis, painting facilities, etc.) and analysis of liquid phase solvent emissions (e.g. petroleum analysis, synthetic lubricants, etc.) and analysis of liquid phase solvent emissions in product gases (e.g. solvents).

INTRODUCTION:

This G.C. procedure employs the polar stationary phase 1, 2, 3-tris (7-cyanoheptyl) propane (TCPP). This material was first employed by H. Mathe et al. for aromatic hydrocarbons, alcohols, ethers, amines and various analyses, while this application permits the simultaneous analysis of aliphatic and aromatic hydrocarbons in admixture.

EQUIPMENT:

GOW-MAC Series 740 Flow Injection Gas Chromatograph
GOW-MAC Model 740-700 Flow Injection Gas Chromatograph
Stationary phase column: 10' x 1/8" packed with 1, 2, 3-tris (7-cyanoheptyl) propane (TCPP) 1, 2, 3-tris (7-cyanoheptyl) propane (TCPP)
Flow: 10-100 ml/min
Microinjection: 0.5 microliter capacity, 500 psi

OPERATING CONDITIONS:

Flow Rate: Carrier: Helium 20 ml/min (for nitrogen quantities: 10 ml/min for nitrogen quantities)
Hydrogen: 30 ml/min
Air: 100 ml/min
Temperatures: Column: 60°C
Injector Port: 80°C
Detector: 110°C
Attenuation: 8 x 10⁻³ (aromatic quantities) 16 x 10⁻³ (aliphatic quantities)

ANALYSIS:

Chromatogram A displays the separation and response for an eight component mixture of aliphatic and aromatic hydrocarbons containing all major peaks of interest in the components. Chromatogram B displays the separation and response for a similar mixture containing the quantities matched from the chromatogram A for the aliphatic components and B for the aromatic components.

GOW-MAC

U.S.A.
CANADA
U.K.

● Four new data sheets from Gow-Mac cover novel topics. One discusses "Forensic and Analytical Toxicological Drug Analysis", the second discusses other "Analytical and Toxicological Drug Analysis", the third discusses "Forensic Toxicology", and the fourth discusses a new instrument, the 740-770P series of gas chromatographs. To receive copies of these data sheets and a list of others available use the Reader Service Card. *Reader Service No. 103.*

● Miles Scientific offers six monoclonal antibodies to four major groups of proteins: collagen Type IV, laminin, keratin sulphate and proteoglycan. (Three antibodies directed to three types of chondroitin sulphate stubs remaining after chondroitinase ABC digestion.) All antibodies are broadly species reactive except collagen Type IV, which appears specific for human tissue and cells only. *Reader Service No. 104.*

● Receptor binding assays are simplified by a new system from Skatron based on the Skatron Cell Harvester. Operation of the instrument involves simultaneous filtration (under reduced pressure) of twelve incubation samples suitably held in Macrowell tubes. A specially designed suction unit head effectively washes the contents of the tubes onto a suitable filter paper — Skatron recommends glass-fibre filter catalogue no. 7034. The Skatron Filterpunch transfers the disks of filter material, six at a time, directly into scintillation vials. *Reader Service No. 105.*

© 1985 Nature Publishing Group

● Now available from Genzyme Corporation for use in glycoprotein research is *N*-glycanase (peptide *N*-glycosidase F). The ultrapure enzyme cleaves all asparagine-linked glycan chains from glycoproteins to yield a carbohydrate-free polypeptide and oligosaccharides with the reducing end of their chitobiose cores intact. The enzyme preparation is also completely free of proteases and other endo- or exoglycosidase activities. The purity and broad specificity for complex, hybrid and high-mannose oligosaccharide structures makes the enzyme useful for structural and functional studies of the sugar and polypeptide portions of glycoproteins. *Reader Service No. 106.*

● A new Biorad cation exchange HPLC column offers high resolution, high sample loads and high recoveries of biologically active proteins and peptides. Called Bio-Gel TSK SP-5PW, the new resin-based column can operate in the pH range 2-12. Either acid (0.1M HCl) or base (0.1M NaOH) washes may be used to clean the column. Pore size is 1,000Å, so the SP-5PW matrix can separate proteins with molecular weights as high as one million with no loss of capacity, resolution or column life. Loading capacity is in the range 1-5 mg for the protein of interest. *Reader Service No. 107.*



The Reichert-Jung Polyvar with new condenser.

● A new condenser for the Reichert-Jung Polyvar photomicroscope overcomes the difficulties inherent in illuminating very large fields at different magnifications. Instead of having to re-arrange the optical elements in the condenser and adjust the aperture and field size to accommodate the objectives being used, the operator can utilize the widefield condenser 0.90/0.32 NA system to provide uniform illumination within a broad magnification range. *Reader Service No. 108.*

ANNOUNCEMENTS

May meetings

14-16 May 1985, **International Conference and Exhibition on Biotechnology**, Boston, USA (B. Burris, Cahners Exposition Group, 7315 Wisconsin Avenue, PO Box 70007, Washington, DC 20088, USA).

14-17 May 1985, **International Society for Neurochemistry Symposium on Excitatory Amino Acids**, Bournemouth, England (Dr P.J. Roberts, Department of Physiology and Pharmacology, University of Southampton, Southampton SO9 3TU, UK).

21-24 May 1985, **Conference on Lasers and Electro-optics**, Maryland, USA (Optical Society of America, 1816 Jefferson Place, N.W. Washington, DC 20036, USA).

22-24 May 1985, **International Congress on Tissue Integration in Oral and Maxillo-Facial Reconstruction**, Brussels, Belgium (Prof. D. van Steenberghe, Dept. of Periodontology, Faculty of Medicine, Catholic University Leuven, Kapucijnenvoer 7, B-3000 Leuven, Belgium).

24-25 May 1985, **International Meeting on Genetic Polymorphism and Fertility in Mammals**, Rome, Italy (Prof. G. Montalenti, Accademia dei Lincei).

June meetings

3-6 June 1985, **Karlsruhe International Conference on Analytical Chemistry in Nuclear Technology**, Karlsruhe, FRG (Prof. Dr Hans J. Ache, Kernforschungszentrum Karlsruhe GmbH, Institut für Radiochemie, Postfach 3640, D-7500 Karlsruhe 1, FRG).

4-6 June 1985, **Symposium on Artificial Intelligence**, COGNITIVA 85, Paris, France (Marie-France Chicane, CESTA, 1 rue Descartes 75005 Paris, France).

6-8 June 1985, **International Conference on Biological Reactive Intermediates**, University of Maryland, USA (Dr R. Snyder, Dept. of Pharmacology and Toxicology, College of Pharmacy, Rutgers University, Piscataway, NJ 08854, USA).

10-14 June 1985, **International Conference on Computers in Chemical Research and Education**, Oberbayern, Germany (Dr J. Brandt, Institut für Organische Chemie, Technische Universität München, Lichtenbergstr. 4, D-8046 Garching, FRG).

14 June 1985, **International Seminar on Immunopathology of Rheumatic Diseases**, Montpellier, France (Prof. J. Sany, Service d'Immuno-Rhumatologie et Readaption Fonctionnelle, Centre Gue-de-Chauliac, Hôpital Saint-Eloi, 34059 Montpellier, France).

17-20 June 1985, **International Conference on Developments in Analytical methods in the Pharmaceutical and Biomedical Sciences**, Camerino, Italy (Prof. I. Antonini, University of Camerino, 62032 Camerino, Italy).

19-21 June, **International Symposium on Molecular Aspects of Neurobiology**,

Florence, Italy (Dr E. Folco, Fondazione Giovanni Lorenzi, Via Montenapoleone, 23-20121 Milan, Italy).

23-27 June 1985, **International Conference on Modern Trends in Activation Analysis**, Copenhagen, Denmark (Dr K. Heydom, Riso National Laboratory, Postbox 49, Dk-4000, Roskilde, Denmark).

25-28 June 1985, **International Conference on Surface Engineering**, Brighton, UK (The Conference Office, The Welding Institute, Abington Hall, Abington, Cambridge CB1 6AL, UK).

26-28 June 1985, **International Congress on Lasers in Medicine and Surgery**, Bologna, Italy (Medicina Viva, Viale dei Mille, 140 43100 Parma, Italy).

July meetings

30 June-5 July, **International Conference on Proteins of Iron Metabolism**, Villeneuve d'Ascq, France (Prof. G. Spik, Université des Sciences et Techniques de Lille I, Laboratoire de Chimie Biologique, F-59665 Villeneuve d'Ascq Cedex France).

1-3 July 1985, **International Symposium on Physico-Chemical Properties and their Role in Environmental Hazard Assessment**, Canterbury, UK (H.L. Bennister, 174 Beverley Drive, Edgeware, Middlesex HA8 5ND, UK).

2-5 July 1985, **International Conference on Improving University Teaching**, Utrecht, The Netherlands (T.B. Massey, The University of Maryland University College, College Park, MD 20742, USA).

7-11 July 1985, **International Symposium for Comparative Research on Leukemia and Related Diseases**, Hamburg, Germany (Prof. F. Deinhardt, Max v. Pettenkofer Institut, Pettenkoferstr. 9a, D-8000, Muenchen 2, FRG).

7-12 July 1985, **International Conference on Natural Products**, North Carolina, USA (Prof. D.G. Kingston, Dept. of Chemistry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA).

7-12 July 1985, **International Conference on Ion Beam Analysis**, Berlin, FRG (IBA'85, HMI Berlin, Postfach 39 01 28, D-1000 Berlin 39, FRG).

7-10 July 1985, **International Congress on Engineering and Food**, Alberta, Canada (ICEF-4 Dept. of Food Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2P6).

8-12 July 1985, **International Meeting on NMR Spectroscopy**, Cambridge, UK (Dr J. Gibson, Secretary (Scientific), The Royal Society of Chemistry, Burlington House, London W1V 0BN, UK).

8-12 July 1985, **International Symposium on Controlled Release of Bioactive Materials**, Geneva, Switzerland (USA: CRS Administrative Office, 16 Not-

tingham Drive, Lincolnshire, IL 60015 312/945-1616. Europe: Symporg S.A. P.O. Box 76, CH-1222, Vesenaz-Geneva, Switzerland 022-524-104).

August meetings

3-8 August 1985, **International Leukocyte Culture Conference** to be held jointly with the National Meeting of the Reticuloendothelial Society, Ithaca, New York, USA. Awards of up to \$500 will be made for papers, on the basis of scientific merit. There will also be an award of \$500 for the winning presentation (RES/LCC Conference Office, Dr Sherwood M. Reichard, Medical College of Georgia, Augusta, Georgia 30912, USA).

11-17 August 1985, **International Anatomical Congress and Exhibition**, Barbican Centre, London, UK (IAC Secretariat, 100 Park Road, London NW1 4RN, UK).

12-14 August 1985, **Symposium on the Ethics of Animal Experimentation**, Stockholm, Sweden (M. Thelestam, Dept. of Bacteriology, Karolinska Institutet, S-104 01 Stockholm, Sweden).

18-21 August 1985, **International Symposium on Order in the Amorphous State of Polymers**, Midland, Michigan, USA (S. Butler, Symposium Secretary, MMI, 1910 West St. Andrews Road, Midland, Michigan 48640, USA).

21-24 August 1985, **International Meeting on Metallothionein and other Low Molecular Weight Metal-binding Proteins**, Zurich, Switzerland (Prof. J.H.R. Kagi, Institute of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland).

25-30 August 1985, **IUPAP Nuclear Physics Conference**, Harrogate, UK (The Institute of Physics, 47 Belgrave Square, London SW1X 8QX, UK).

26-30 August 1985, **International Symposium on Nuclear Medicine and Related Medical Applications of Nuclear Techniques in Developing Countries**, Vienna, Austria (Int. Atomic Energy Agency, IAEA-SM-283, Vienna International Centre, PO Box 100 a 1400 Vienna, Austria).

29-31 August 1985, **20th Meeting of the European Association for the Study of the Liver**, Helsinki, Finland (EASL, The Finnish Medical Society Duodecim, Mrs Leena Antikainen, Runeberginkatu 47 A, SF-00260 Helsinki, Finland).

September meetings

2-6 September 1985, **International Symposium on Marine Natural Products**, Paris, France (Drs A. Ahond and C. Poupat, ICSN du CNRS, 91190 Gif s/Yvette France).

2-7 September 1985, **European Drosophila Research Conference**, Balatonszeplak, Hungary (J. Szabad, Institute of Genetics, Biological Research Center, H-6701 Szeged, Hungary).

3-5 September 1985, **Symposium of the**

Please send 'Announcements' entries to:
The Editor, Nature,
4 Little Essex Street,
London WC2R 3LF, UK

Uranium Institute, London, UK (Symposium Secretariat, Conference Associates UIS, 27A Medway Street, London SW1P 2BD, UK).

3-6 September 1985, **Noordwijkerhout Symposium on Innovative Approaches in Drug Research**, Noordwijkerhout, The Netherlands (Dr J.F. Rodrigues de Miranda, c/o Prof. Dr H. Timmerman, Vrije Universiteit, Department of Pharmacology, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands).

3-6 September 1985, **European Nuclear Medicine Congress**, London, UK (European Nuclear Medicine Congress 1985, Institute of Nuclear Medicine, Middlesex Hospital Medical School, Mortimer Street, London W1N 8AA, UK).

8-13 September 1985, **International Congress of Pure and Applied Chemistry**, Manchester, UK (Dr J.F. Gibson 30th IUPAC Congress, Royal Society of Chemistry, Burlington House, London W1V 0BN, UK).

8-13 September 1985, **European Immunology meeting**, Jerusalem, Israel (Prof. M. Schlesinger, Secretariat of the 7th Immunology Meeting, PO Box 983, Jerusalem 91009, Israel).

9-13 September 1985, **IUPAC International Congress of Pure and Applied Chemistry**, Manchester, UK (Dr J. Gibson, 30th IUPAC Congress, Royal Society of Chemistry, Burlington House, London W1V 0BN, UK).

3 Sept-4 October 1985, **International Remote Sensing Workshop**, South Dakota, (US Geological Survey, 917 National Center, Reston, VA 22092, USA).

8-13 September 1985, **International Conference on Biothermodynamics**, Styria, Austria (Dr P. Laggner, Institut für Röntgenfeinstrukturforschung, der Österreichischen Akademie der Wissenschaften und des Forschungszentrums, Graz, Austria).

8-13 September 1985, **Laurentian Hormone Conference**, Alberta, Canada (The Laurentian Hormone Conference, 45 Shattuck Street, Boston, MA 02115, USA).

9-10 September 1985, **International Symposium on Drug Analysis: Current Challenges**, Ottawa, Canada (Dr G.L. Mattok, Bureau of Drug Research, Health Protection Branch, Tunney's Pasture, Ottawa, Ontario, Canada K1A 0L2).

9-14 September 1985, **International Conference on Magnetospheric Systems**, Toulouse, France (Centre National d'Études Spatiales, Département des Affaires Universitaires, 18 avenue Édouard-Belin, 31055 Toulouse, France).

Appointments

Dr Archie Johnston (director of the Health and Safety Executive's Research and Laboratory Services Division) has been appointed to the Health and Safety Executive to make up the three man team. The executive is responsible for implementing the

provisions of the health and Safety at Work Act 1984.

Professor John L. Jinks has been appointed to be Secretary and Deputy Chairman to the Agricultural and Research Council from 1 May 1985. He has been chairman of the council's plants and soils research committee for the past two years. **Dr Jo Eirik Asvall** has been appointed Regional Director for Europe of the World Health Organization (WHO). For the past six years Dr Asvall has been Director of Programme Management at the WHO Regional Office for Europe.

Sir Francis Tombs is to become chairman of The Engineering Council on 1 May 1985 for a three-year period. Sir Francis is a past president of the Institution of Electrical Engineers.

Mr John Webb (managing director of Bayer Limited) has been appointed Chairman of the Chemical Industry Safety, Health and Environment Council (CISHEC), the body responsible for the UK chemical industry's position on health, safety and the environment.

Dr Ronald S. Paul has been elected Chief Executive Officer of the Batelle Memorial Institute having been president and Chief Operating Officer for the past three years.

Awards

Dr D.A. Henderson, who led the global campaign that successfully eradicated smallpox, has been awarded the Albert Schweitzer International Prize for medicine by the Burroughs Wellcome Fund (supported by the Burroughs Wellcome Company). Dr Henderson is dean of The John Hopkins School of Public Health, Baltimore, Maryland.

Loughborough University is to confer the honorary degree of Doctor of Technology on Olympic gold medallist and world record breaking athlete **Sebastian Coe**. Coe has a BSc degree in Economics from Loughborough, where he was an undergraduate, research student and research assistant. It is hoped that the ceremony will take place in July this year.

750 Overseas Research Students (ORS) Awards will be offered on a competitive basis in 1985 to overseas postgraduate students. The scheme is administered by the Committee of Vice-chancellors and Principals of the universities of the United Kingdom. Each award will cover the difference between the tuition fee for a home postgraduate student and the "full-cost" fee chargeable to an overseas postgraduate student. Candidates must be overseas graduate students who, in session 1985-86, are commencing full time study as registered research students for a higher degree at one of the academic institutions participating, or are already on such a course and do not have an ORS award. Awards may be held in any field of study by students of outstanding merit and research

potential. Further details and application forms from the registrar of any of the academic institutions taking part in the scheme, including Aberdeen, Leeds and Sussex Universities.

The Simon Fraser University, Canada, announces the R.H. Wright award in Olfactory Research of \$25,000. The award will be made to an individual who has recorded outstanding achievements in research in olfaction. Names of proposed candidates and supporting information should be sent to B.F. Clayman, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada.

The European Science Foundation is offering short and long term research fellowships in toxicology. Fellowships can be awarded to applicants of any nationality working in any field of the natural and life sciences, provided that their project contributes to furthering research in toxicology within Europe. The deadline for submitting completed applications is 30 September 1985, and application forms may be obtained from: ESF Research Fellowships in Toxicology (PGT), Dr John Smith, European Space Foundation, 1 quai Lezay-Marnesia - F - 67000 Strasbourg, France.

The British Ecological Society is offering grants ranging from £50 to £500 to amateur and professional scientists undertaking ecological research and survey, usually within the United Kingdom, but grants are also given to projects abroad. Application forms are available from the British Ecological Society (S.E.P.G.), Burlington House, Piccadilly, London W1V 0LQ. Further details from Dr M.B. Usher, Department of Biology, University of York, York YO1 5DD.

The Gerontological Society of America invites nominations for its 1985 Brookdale Awards for distinguished contributions to gerontology. Three awards of a \$25,000 cash prize will be given to gerontologists recognized nationally and internationally in their field, two of whom will be US citizens and one a non-US citizen. Further details from Linda Harootyan, The Gerontological Society of America, 1411K St, NW, Suite 300, Washington, DC 20005.

Miscellaneous

Columbia University chemical engineer **Helmut W. Schulz**, at 72, has won official recognition for a 44-year old idea that was long wrapped in military secrecy. The US government has awarded him \$100,000 in recognition of an innovative proposal he made as a young man for an energy efficient gas centrifuge method of uranium enrichment. This is the largest sum ever granted for an invention put under military wraps and thus unavailable for use by its inventor. □